

# **ANÁLISIS DE MÉTODOS DE PROCESAMIENTO DE IMÁGENES ESTEREOSCÓPICAS FORESTALES**

**MÁSTER EN INVESTIGACIÓN  
EN INFORMÁTICA  
2007-08**

**PROYECTO FIN DE MÁSTER**



Departamento de Ingeniería del Software e Inteligencia Artificial  
Facultad de Informática  
Universidad Complutense  
MADRID

Alumno:  
Carlos Zafra Fernández de Tejada

Director:  
Gonzalo Pajares Martinsanz



# Resumen

Mediante imágenes de zonas boscosas se puede determinar la presencia de ciertas regiones a través de la segmentación de las mismas así como detectar la presencia de ciertas componentes (hojas, troncos, cielo) a través de técnicas como la extracción de bordes y la clasificación de texturas combinadas con la segmentación de regiones. El objetivo de este trabajo se centra en la búsqueda de un algoritmo que sea capaz de obtener a partir de un sistema estéreo, los troncos de los árboles presentes en la imagen, para su posterior procesamiento (calcular altura y volumen) y con ello la cantidad de madera presente en la zona objeto de interés. En general este tipo de técnicas comienzan a ser muy estudiadas, porque permiten automatizar tareas con el consiguiente ahorro de tiempo.

**Palabras clave:** imágenes forestales, clasificación, texturas, bordes, estéreo visión.

# Agradecimientos

A Fernando Pita Ingeniero de Montes y a la Dra. Isabel Cañellas quienes a través del Centro de Investigación Forestal (CIFOR) del Instituto Nacional de técnicas Agroalimentaria (INIA) han facilitado las imágenes estereoscópicas con las que se ha realizado este trabajo.



# Índice

<b>ÍNDICE.....</b>	<b>III</b>
<b>1. INTRODUCCIÓN.....</b>	<b>5</b>
1.1. IDENTIFICACIÓN DEL PROBLEMA .....	5
1.2. MOTIVACIÓN Y OBJETIVOS DE INVESTIGACIÓN .....	7
1.2.1. Motivación.....	7
1.2.2. Objetivos generales .....	8
1.2.3. Objetivos específicos.....	8
1.3. METODOLOGÍA HISTÓRICA .....	8
1.4. APORTACIONES A LA INVESTIGACIÓN .....	11
<b>2. REVISIÓN DE MÉTODOS .....</b>	<b>13</b>
2.1. INTRODUCCIÓN.....	13
2.2. MÉTODOS DE CLASIFICACIÓN DE TEXTURAS NATURALES .....	13
2.3. DESCRIPCIÓN DE LOS CLASIFICADORES SELECCIONADOS.....	15
2.3.1. Agrupamiento borroso.....	15
2.3.2. Clasificador Bayesiano.....	18
2.3.3. Algoritmo Generalizado de Lloyd.....	21
2.4. MÉTODOS LINEALES DE EXTRACCIÓN DE BORDES .....	23
2.5. OPERADORES DE EXTRACCIÓN DE BORDES .....	24
2.5.1. Roberts.....	24
2.5.2. Prewitt .....	24
2.5.3. Sobel.....	25
2.6. DESCRIPTORES ESTADÍSTICOS DE TEXTURA .....	25
2.6.1. Media y desviación estándar .....	25
2.6.2. Entropía .....	26
2.7. IDENTIFICACIÓN DE LÍNEAS RECTAS: TRANSFORMADA DE HOUGH .....	26
<b>3. DISEÑO DE ALGORITMOS .....</b>	<b>31</b>
3.1. INTRODUCCIÓN.....	31
3.2. IDENTIFICACIÓN DEL PROBLEMA .....	33
3.3. ANÁLISIS DE LAS IMÁGENES: PROBLEMÁTICA.....	34
3.4. DISEÑO DEL MÉTODO MEDIANTE TÉCNICAS DE CLASIFICACIÓN .....	36
3.5. DISEÑO DEL MÉTODO USANDO UN FILTRO DE ENTROPÍA.....	42
3.6. DISEÑO DE UN MÉTODO MEDIANTE EXTRACCIÓN DE BORDES .....	46
<b>4. ANÁLISIS DE RESULTADOS .....</b>	<b>51</b>
4.1. OBJETIVOS DEL ANÁLISIS.....	51
4.2. DESCRIPCIÓN DE LAS IMÁGENES UTILIZADAS .....	51
4.3. ANÁLISIS DE RESULTADOS.....	53
4.3.1. Criterio de validación .....	53
<b>5. CONCLUSIONES Y TRABAJO FUTURO.....</b>	<b>57</b>
5.1. SOBRE RESULTADOS .....	57
5.2. CUMPLIMIENTO DE LOS OBJETIVOS PROPUESTOS .....	58
5.3. TRABAJOS FUTUROS .....	58
<b>6. BIBLIOGRAFÍA.....</b>	<b>59</b>



# Capítulo 1

## 1. Introducción

### 1.1. Identificación del problema

Es preciso reconocer que hoy por hoy la visión por computador a veces no es la mejor solución a un problema. Existen muchas ocasiones en las que el problema es tan complejo que la solución humana es lo mejor. Por ejemplo imaginemos una conducción de un vehículo en una carretera con tráfico intenso. Pero a veces, las soluciones humanas tienden a ser inexactas o subjetivas y en ocasiones lentas y presentan una ausencia de rigor así como una pobre percepción (Marshall y Martin 1993). No obstante, la solución humana es menos estructurada que la solución artificial y muchos problemas de visión por computador requieren un nivel de inteligencia mucho mayor que el que la máquina pueda ofrecer. El sistema de visión humana puede describir automáticamente una textura en detalle, un borde, un color, una representación bidimensional de una tridimensional, ya que puede diferenciar entre imágenes de diferentes personas, firmas, colores, etc., puede vigilar ciertas zonas, diagnosticar enfermedades a partir de radiografías, etc. Sin embargo, aunque algunas de estas tareas pueden llevarse a cabo mediante visión artificial, el software o el hardware necesario no consigue los resultados que serían deseables.

Aun a pesar de las limitaciones expuestas, cada día es mayor el número de aplicaciones de la visión artificial. Por ello sólo es posible dar una pequeña pincelada sobre las múltiples aplicaciones en las que la visión artificial se ha aplicado hasta el momento. Aquí solamente proporcionamos algunas ideas para generar aplicaciones en la mente del lector. Obviamente los ejemplos que se presentan son aplicables a cualquier otro proceso diferente y en otro campo industrial o científico también diferentes. En Newman y Jain (1995) o Davies (1997) se puede encontrar una amplia revisión sobre la inspección visual y aplicaciones de la visión a diferentes campos de la vida real, que nos dan idea del amplio abanico de posibilidades. Para dar idea tan sólo algunas de tales aplicaciones:

- 1) Control de cultivos en agricultura, propiciado por la necesidad de conocer los cultivos programados para la recepción de subvenciones u otros aspectos relacionados.
- 2) Cómputo y medición de parcelas agrícolas y tipo de cultivos a los que se dedica.
- 3) Control de riegos agrícolas
- 4) Agricultura de precisión para aplicar herbicida en el tratamiento de malas hierbas de forma selectiva evitando la contaminación medioambiental y la reducción de costes de producción.
- 5) Evaluación de catástrofes naturales: fuegos, daños por inundaciones, heladas en cultivos agrícolas, nevadas, etc.
- 6) Detección de cambios en determinadas zonas, principalmente urbanas para el control de edificaciones o impactos medioambientales
- 7) Vigilancia en prevención de catástrofes, por ejemplo fuegos o inundaciones
- 8) Control de fenómenos meteorológicos como es el retroceso de determinadas playas
- 9) Vigilancia: forestal, marítima.
- 10) Detección de infraestructuras: carreteras, caminos forestales, cañadas reales, etc.
- 11) Finalmente, destacamos la aplicación objeto de esta memoria, que en sí misma constituye una aplicación novedosa al tratarse de la obtención del volumen de madera en una zona boscosa.

Sobre esta última aplicación y hasta donde se ha podido investigar, no existen trabajos previos específicos que permitan establecer un punto de partida concreto.

En general la mayoría de herramientas disponibles implementan algunos de los métodos clásicos de clasificación, siendo necesaria la intervención del usuario mediante programación para abordar ciertas tareas que involucran aspectos de clasificación de texturas, lo cual no siempre es factible.

Además, y lo que es más importante, en muchos casos no existe la posibilidad de llevar a cabo la investigación necesaria para abordar la problemática, particularmente cuando los métodos clásicos no producen los resultados esperados.



La presente memoria se organiza en cinco capítulos. En lo que resta del presente capítulo se exponen los objetivos que se plantearon en la investigación aquí recogida, en base a los planteamientos previos y que tratan de dar solución a la problemática existente. También se exponen las estrategias realizadas para abordarlos. En el capítulo 2 se realiza una revisión del estado del arte en el tema de clasificación de texturas. En el capítulo 3 se propone la estrategia de solución a la problemática planteada, junto con los pasos del proceso que la conforman. En el capítulo 4 se diseña una estrategia de pruebas con el objetivo de verificar la eficacia de la solución propuesta cuando se compara con otras estrategias existentes. Finalmente en el capítulo 5 se extraen las conclusiones pertinentes y se exponen las líneas de investigación futuras, que dan pie al inicio de la investigación en los estudios de Doctorado.

## **1.2. Motivación y objetivos de investigación**

### **1.2.1. Motivación**

Como se ha mencionado anteriormente, la investigación consiste en la identificación de las texturas existentes en zonas boscosas con el fin último de estimar el volumen de madera existente a partir de un sistema de Visión estereoscópica.

Este trabajo tiene su origen en el proyecto de investigación titulado “Análisis estereoscópico y reconocimiento de texturas en imágenes forestales obtenidas mediante dispositivo de medición MU-200501738”. Dicho proyecto se desarrolla bajo un contrato de investigación al amparo del artículo 83 de la LOU con referencia INIA-445/2007, firmado con el Instituto Nacional de Investigación Agraria y Agroalimentaria (INIA) en colaboración también con la Escuela Técnica Superior de Ingenieros de Montes (ETSIM) de la Universidad Politécnica de Madrid siendo investigador principal del mismo el director de este trabajo de investigación.

La captura de las imágenes se lleva a cabo en el bosque objeto de análisis, situándose en diferentes áreas hasta completar la superficie total bajo análisis. Las imágenes se procesan manualmente en los laboratorios de la ETSIM o del INIA para estimar el volumen de madera, siendo el objeto de este trabajo la investigación de técnicas con el objetivo de automatizar el proceso de estima.

### **1.2.2. Objetivos generales**

- 1) Aprender a manejar referencias bibliográficas, así como la forma de abordar las investigaciones
- 2) Aprender y familiarizarse con la mayoría de técnicas existentes de visión por computador.
- 3) Analizar el problema y ver cuales de los métodos anteriores o combinaciones de los mismos podrían aplicarse a la resolución del problema de la estima automática de la madera existente en cada par de imágenes estereoscópicas.
- 4) Analizar las ventajas e inconvenientes de cada uno de ellos así como su problemática
- 5) Dar solución a un problema realizando un aporte de carácter investigador.

### **1.2.3. Objetivos específicos**

- 1) Identificar y diseñar un algoritmo mediante el cual se pueda dar una solución lo más precisa posible al problema.
- 2) Identificar una posible mejora de los métodos existentes o dar una solución novedosa a la problemática de la clasificación de las texturas naturales.
- 3) Implementar la mejora propuesta con el fin de analizar su comportamiento en base a los resultados obtenidos.
- 4) Identificar líneas de investigación futuras
- 5) Difundir los resultados mediante publicaciones científicas u otro tipo de divulgación

## **1.3. Metodología histórica**

A continuación se expone de forma cronológica los pasos seguidos para llevar a cabo la investigación junto con la actividad desarrollada en cada uno de ellos.

Se inicia el trabajo de investigación en el marco del proyecto mencionado previamente, donde la clasificación de texturas aparece como un reto para dar respuesta a la necesidad de métodos automáticos de medición de árboles.

Ya dentro del programa Master, y en base a la problemática identificada a la que hay que proporcionar una solución, se identifican algunos de los cursos, que en principio parecen más acordes con el tema objeto de investigación. En percepción Visual se aborda el tratamiento de las imágenes digitales desde el punto de vista del reconocimiento y el aprendizaje. Comienzan a sentarse las bases de la investigación desde el inicio del curso, que culmina con un trabajo sobre reconocimiento de patrones en el ámbito de las texturas naturales.

Se estudian los métodos más avanzados de clasificación de texturas que podrían ser viables a priori. Algunos ejemplos de clasificadores que se estudiaron fueron el *Clasificador paramétrico Bayesiano*, el *algoritmo de Lloyd* y el método de *Agrupamiento borroso o Fuzzy clustering*.

Sin embargo de entre todos los posibles métodos, se opta por utilizar una forma híbrida de los mismos conocida como *votador mayoritario (majority voting)* (Kittler y col.,1998); Kuncheva, 2004; Grim y col. 2002; Martinez-Otzeta, 2006). Las técnicas de clasificación híbridas han sido estudiadas en el trabajo de Guijarro (2007) con un cierto éxito garantizado. Se han realizado diferentes pruebas para verificar la eficacia de los métodos de clasificación..

La clasificación de texturas en sí misma no resulta suficiente para resolver el problema planteado. Por este motivo, se optó por intentar buscar una solución en el campo de la extracción de bordes y la segmentación. De nuevo se experimentó con diversos métodos existentes en la literatura. Algunos ejemplos de operadores de extracción de bordes fueron:

- Sobel
- Prewitt
- Roberts
- Laplaciano

De entre todos los que se probaron el filtro de primer orden de Sobel fue con el que se obtuvieron los resultados más prometedores. Sin embargo, este método necesitaba de una

ayuda extra para poder trabajar con él. Era necesario aplicar algunos operadores adicionales para obtener resultados similares a los obtenidos mediante clasificación. Tras perfeccionar el proceso se obtuvieron resultados muy similares a los obtenidos mediante un clasificador, pero con un tiempo de procesamiento bastante menor.

El siguiente paso consistió en estudiar el color de las imágenes. Ningún método basado en las propiedades del color o en las transformaciones de dominio como la HSV dio un resultado mínimo aceptable. Esto lógicamente se debe a que todas las imágenes de las que disponíamos estaban comprimidas en formato JPEG, por lo que la compresión de este formato produce efectos no deseados en el color, uno de ellos es la adquisición de tonalidad azul en hojas de árboles que en realidad son verdes. Este hecho resulta determinante debido a que una de las exigencias del proceso consiste en la distinción de las hojas de los árboles frente a los troncos y ramas, que son los generadores de la madera.

El último intento de mejorar, el procesamiento inicial de la imagen consistió en estudiar algunas de las medidas estadísticas de la imagen como son la desviación estándar o la entropía. De nuevo se observó que este tipo de métodos (especialmente la entropía) producían resultados tan buenos como los clasificadores. Estos métodos se encuadran específicamente dentro de los denominados descriptores de texturas.

A continuación se trató de encontrar un método que eliminara el “ruido” de los resultados obtenidos mediante los clasificadores de texturas y de extracción de bordes. Finalmente se encontró un método que conseguía “limpiar” las imágenes de cara al objetivo de identificación de la madera.

Finalmente, dado que las especies arbóreas de los bosques donde se utiliza el sistema, son en su mayoría especies con troncos rectos o aproximadamente rectos, se han estudiado métodos de identificación de líneas rectas como es el caso de la transformada de Hough (1962).

Queda todavía por resolver un aspecto relacionado con el procedimiento de correspondencia entre los pares del sistema estéreo. Sin embargo, el estudio de este

problema queda fuera del alcance del presente trabajo debido al volumen de investigación que conlleva por sí mismo.

Los resultados, aunque en este momento de la investigación no son los ideales, sí que se presentan ciertamente como prometedores. Seguramente, con unas imágenes de mayor calidad en las que se pueda usar la información del color se puedan desarrollar métodos más precisos. Hay que recordar que este campo se encuentra todavía en una etapa inicial, y aún quedan muchas ideas y métodos que desarrollar.

## **1.4. Aportaciones a la investigación**

Como resumen de todo lo anterior las aportaciones de investigación realizadas en el presente trabajo son las siguientes:

- 1) Diseño de métodos de detección e identificación de los troncos de los árboles
- 2) Estudiar la precisión y bondad de los resultados obtenidos.
- 3) Sentar las bases de trabajos futuros, y proponer futuras vías de investigación en base a los resultados obtenidos.



# Capítulo 2

## 2. Revisión de métodos

### 2.1. Introducción

Siguiendo los pasos de investigación descritos en la sección 1.3 en este capítulo se describen los métodos investigados en relación a:

- Métodos de clasificación
- Extractores de bordes
- Descriptores estadísticos de texturas
- Identificación de líneas rectas: transformada de Hough

### 2.2. Métodos de clasificación de texturas naturales

Se han propuesto diferentes métodos de clasificación de texturas naturales en la literatura, entre los que destacan algunos tales como: el clasificador Bayesiano, Vecino más próximo (K-Nearest), Redes Neuronales (Pajares y col. 2001; Gonzalez y Woods, 2002) o Cuantización vectorial (Learning Vector Quantization) (Giacinto y col. 2000, Chan y col. 2003, del Frate 2007), Agrupamiento borroso (Pajares y Cruz, 2002, Pajares y col. 2002).

La conducta de diferentes características ha sido estudiada en clasificación de texturas. Se entiende por característica alguna propiedad que describa el contenido de la textura (Valdovinos y col. 2005, Puig y García, 2006 o Hanmandlu y col. 2004). Existen dos orientaciones importantes en este aspecto, a saber: *a nivel de píxeles y a nivel de regiones*.

A nivel de píxeles se considera cada píxel de forma individual y se trata de clasificar el mismo como perteneciente a una determinada clase. En este caso las propiedades que se evalúan son sus componentes espectrales, esto es el color (Rud y col. 2006, Maillard, 2003). En cuanto a los modelos de color utilizados destaca el modelo RGB, no obstante

cualquiera de los modelos de color (HSV, YIQ, etc.) puede ser utilizado. En el trabajo de Yu y col. (2003) se analizan algunos de ellos. También en Kumar y col. (2001, 2002) se estudian algunos aspectos relacionados con el análisis multispectral.

Los métodos a nivel de regiones no sólo tienen en cuenta cada píxel de forma individual sino los vecinos que le rodean. En este sentido lo que se evalúa son determinadas formas de aparición de las texturas siguiendo ciertos patrones. Existen diversas bases de imágenes con este tipo de texturas, figura 2.1 tomada de Brodatz (2008). Existen otras bases de imágenes de textura tales como BME textura data base (2008) o VisTex (2008).



Figura 2.1 Diferentes ejemplos de texturas

En el conjunto de imágenes de prueba del que disponemos, hallamos las siguientes características:

- 1) En la figura 2.2 no aparecen patrones de textura repetitivos como en la figura 2.1, que permitan aplicar los descriptores que se usan para su tratamiento
- 2) No existe una única textura.
- 3) Existe una clara distinción en las bandas espectrales (colores). Aunque debido a la compresión JPEG de las imágenes esta información no es fiable.





Figura 2.2 Ejemplo de imagen analizada

Como consecuencia de las consideraciones expuestas anteriormente y aún siendo conscientes de que alguna de ellas ya se ha mencionado, en el presente trabajo de investigación se ha optado por lo siguiente:

- Utilizar una estrategia a *nivel de píxel*
- Seleccionar varios clasificadores clásicos, que han sido ampliamente probados y verificados en la literatura aunque no se hayan utilizado de forma específica en el tratamiento de las texturas naturales.

En principio todos los métodos aquí recogidos dan resultados muy similares, es por ello que elegiremos el que sea más fácil de implementar. En nuestro caso será el algoritmo generalizado de Lloyd que se describe en la siguiente sección.

## 2.3. Descripción de los clasificadores seleccionados

### 2.3.1. Agrupamiento borroso

El objetivo de la técnica de agrupamiento conocida como *Agrupamiento borroso o Fuzzy Clustering* consiste en dividir  $n$  objetos  $x \in X$  caracterizados por  $p$  propiedades en  $c$  clústeres o grupos. Supongamos el conjunto de datos  $X = \{x_1, x_2, \dots, x_N\} \in \mathbb{R}^p$  un

subconjunto del espacio real  $p$ -dimensional  $\mathfrak{R}^p$ . Cada  $x_k = \{x_{k_1}, x_{k_2}, \dots, x_{k_p}\} \in \mathfrak{R}^p$  se denomina vector de características,  $x_{k_j}$  es la  $j$ -ésima característica de la observación  $x_k$ .

Este clasificador puede encontrarse perfectamente especificado en las siguientes referencias clásicas: Bezdek (1981), Duda y col., (2000), Pajares y col. (2002) o Zimmerman (1991).

Puesto que los elementos de un clúster deben ser tan similares entre sí como sea posible y a la vez deben ser tan diferentes a los elementos de otros clústeres como también sea posible, el proceso se controla por el uso de medidas de similitud basadas en distancias. Así la similitud o la diferencia entre dos puntos  $x_k$  y  $x_l$  puede interpretarse como la distancia entre esos puntos.

Una distancia entre dos objetos de un universo  $X$  es una función  $d : X \times X \rightarrow \mathbb{R}^+$  que toma valores reales que denotamos  $d(x_k, x_l) = d_{kl} \geq 0$  y que cumple tres propiedades:

1.  $d_{kl} = 0 \Leftrightarrow x_k = x_l$  para todo  $0 \leq k, j \leq N$
2. simetría:  $d_{kl} = d_{lk}$  para todo  $0 \leq k, l \leq N$
3. Desigualdad triangular  $d_{kl} \leq d_{kj} + d_{jl}$  para todo  $0 \leq k, j, l \leq N$  (2.1)

Cada partición del conjunto  $X = \{x_1, x_2, \dots, x_n\}$  puede enfocarse desde dos perspectivas: fuzzy y no fuzzy. Una partición no fuzzy se conoce en terminología inglesa como “crisp”. Si se desea realizar una partición del conjunto  $X$  en  $c$  clústeres tendremos  $S_i \{i = 1, \dots, c\}$  subconjuntos.

Para cada partición,  $S_i$  define un conjunto borroso  $\mu_i$  sobre el universo  $X$ ,  $\mu_i : X \rightarrow [0, 1]$  que asigna lo que se conoce como grado de pertenencia  $\mu_{ik}$  de cada objeto  $x_k$  al subconjunto  $S_i$ . Así pues denotamos dicho grado de pertenencia del elemento  $x_k$  al clúster  $S_i$  como  $\mu_i(x_k) = \mu_{ik}$ . En el caso de conjuntos “crisp” un objeto  $x_k$  se dice que pertenece a un  $S_i$  dado y no pertenece al resto. Esto se expresa con los valores discretos  $\{0, 1\}$  de la siguiente forma  $\mu_{ik} = 1$  para indicar pertenece y  $\mu_{ik} = 0$  para expresar que no pertenece. Por el contrario, en el caso de conjuntos fuzzy se dice que un objeto puede pertenecer a diferentes subconjuntos y así se habla por ejemplo de que  $x_k$  pertenece a un conjunto

$S_i$  con grado de pertenencia  $\mu_{ik}$  y a  $S_j$  con grado de pertenencia  $\mu_{jk}$ . Como ejemplo, supongamos que se tienen tres conjuntos  $S_i$ ,  $S_j$  y  $S_h$ , en este caso podríamos decir que el objeto  $x_k$  pertenece a los conjuntos con los siguientes grados de pertenencia  $\mu_{ik} = 0.4$ ,  $\mu_{jk} = 0.5$  y  $\mu_{hk} = 0.1$ . Los valores tomados pertenecen al intervalo continuo  $[0,1]$ .

Dado  $X = \{x_1, x_2, \dots, x_n\}$  y el conjunto  $V_{cn}$  de todas las matrices reales de dimensión  $c \times n$ , con  $2 \leq c < n$ . Se puede obtener una matriz representando la partición de la siguiente manera  $U = \{\mu_{ik}\} \in V_{cn}$ . Tanto en el supuesto “crisp” como en el fuzzy se deben cumplir las siguientes condiciones:

- 1)  $\mu_{ik} \in \{0,1\}$  *crisp* o  $\mu_{ik} \in [0,1]$  *fuzzy*  $1 \leq i \leq c; 1 \leq k \leq n$
- 2)  $\sum_{i=1}^c \mu_{ik} = 1 \quad 1 \leq k \leq n$  (2.2)
- 3)  $0 < \sum_{k=1}^n \mu_{ik} < n \quad 1 \leq i \leq c$

Para ilustrar los conceptos anteriores sea  $X = \{x_1, x_2, x_3\}$  con él podríamos construir las siguientes particiones teniendo en cuenta que  $c = 2$

“*crisp*”

$$U = \begin{bmatrix} x_1 & x_2 & x_3 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} x_1 & x_2 & x_3 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

“*fuzzy*”

$$U = \begin{bmatrix} x_1 & x_2 & x_3 \\ 0.3 & 0.5 & 0 \\ 0.7 & 0.5 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} x_1 & x_2 & x_3 \\ 0.9 & 0.4 & 0.2 \\ 0.1 & 0.6 & 0.8 \end{bmatrix}$$

La localización de un clúster  $S_i$  se representa por su centro  $v_j = \{v_{j_1}, v_{j_2}, \dots, v_{j_p}\} \in \mathbb{R}^p$  con  $j = 1, \dots, c$ , alrededor del cual se concentran los objetos.

La definición básica de llevar a cabo el problema de la partición fuzzy para  $m > 1$  consiste en minimizar la función objetivo según la ecuación siguiente:

$$\min z_m(U; v) = \sum_{k=1}^n \sum_{j=1}^c \mu_{jk}^m \|x_k - v_j\|_G^2 \quad (2.3)$$

donde  $G$  es una matriz de dimensión  $p \times p$  que es simétrica y definida positiva. Así se puede definir una norma general del tipo,

$$\|x_k - v_j\|_G^2 = (x_k - v_j)^t G (x_k - v_j) \quad (2.4)$$

Diferenciando la función objetivo para  $v_i$  (suponiendo constante  $U$ ) y  $\mu_{jk}$  (suponiendo constante  $v$ ) y aplicando la condición de que  $\sum_{j=1}^c \mu_{jk} = 1$ , se obtiene,

$$v_j = \frac{1}{\sum_{k=1}^n (\mu_{jk})^m} \sum_{k=1}^n (\mu_{jk})^m x_k \quad j = 1, \dots, c \quad (2.5)$$

$$\mu_{jk} = \frac{\left( \frac{1}{\|x_k - v_j\|_G^2} \right)^{2/m-1}}{\sum_{h=1}^c \left( \frac{1}{\|x_k - v_h\|_G^2} \right)^{2/m-1}} \quad j = 1, \dots, c; k = 1, \dots, n \quad (2.6)$$

el exponente  $m$  se conoce como peso exponencial y disminuye la influencia del ruido al obtener los centros de los clústeres, reduciendo la influencia de los valores pequeños de  $\mu_{jk}$  (puntos lejos de  $v_j$ ) frente a valores altos de  $\mu_{jk}$  (puntos cercanos a  $v_j$ ). Cuanto mayor sea  $m > 1$  mayor es dicha influencia.

Existen diferentes medidas escalares para validar la partición, siendo el objetivo encontrar el número de clústeres que obtenga las mejores medidas. De entre todos ellos destaca por su importancia (Balasko y col. 2006, Kim y col. 2003) el *Coeficiente de partición (CP)* que mide la cantidad de solapamiento entre los clústeres

$$CP = \frac{1}{n} \sum_{j=1}^c \sum_{i=1}^n (\mu_{ji})^2 \quad (2.7)$$

Cuanto más próximo sea este valor a la unidad tanto mejor será la clasificación de los datos llegando a ser una partición pura cuando toma el valor de la unidad.

### 2.3.2. Clasificador Bayesiano

En el caso general y más típico de una distribución de probabilidad Gaussiana o Normal multivariable, ni la media  $\mathbf{m}$  ni la matriz de covarianza  $C$  son conocidas. Por tanto, esos parámetros desconocidos constituyen las componentes del vector de parámetros  $\mathbf{w} = \{\mathbf{m}, C\}$  consideremos el supuesto univariable con  $\mathbf{m} = m$  y  $C = \sigma^2$ , en cuyo caso

$$\ln p(x_i/\mathbf{w}) = -\frac{1}{2} \ln 2\pi C - \frac{1}{2C} (x_i - m)^2 \quad (2.8)$$

$$\nabla_{\mathbf{w}} \ln p(x_i/\mathbf{w}) = \begin{bmatrix} \frac{1}{C} (x_i - m) \\ -\frac{1}{2C} + \frac{(x_i - m)^2}{2C^2} \end{bmatrix} \quad (2.9)$$

La minimización sobre los datos de entrenamiento conduce ahora a las condiciones,

$$\frac{1}{n} \sum_{i=1}^n \frac{1}{\hat{C}} (x_i - \hat{m}) = 0 \quad -\frac{1}{n} \sum_{i=1}^n \frac{1}{2\hat{C}} + \frac{1}{n} \sum_{i=1}^n \frac{(x_i - \hat{m})^2}{2\hat{C}^2} = 0 \quad (2.10)$$

donde  $\hat{m}$  y  $\hat{C}$  son las estimas de máxima verosimilitud para  $m$  y  $C$ , respectivamente.

Sustituyendo  $\hat{m}$  y  $\hat{\sigma}^2 = \hat{C}$  obtenemos las estimas de máxima verosimilitud para  $m$  y  $\sigma^2$

$$\hat{m} = \frac{1}{n} \sum_{i=1}^n x_i \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2 \quad (2.11)$$

Aunque el análisis del caso multivariable es básicamente muy similar, se requiere mucha más manipulación. El resultado muy bien conocido en estadística es que las estimas de máxima verosimilitud para  $\mathbf{m}$  y  $C$  están dadas por,

$$\mathbf{m} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \quad C = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^t \quad (2.12)$$

La expresión (2.12) nos dice que la estima de máxima verosimilitud para el vector media es la media simple. La estima de máxima verosimilitud para la matriz de covarianza es la media aritmética de las  $n$  matrices  $(\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^t$ . Puesto que la verdadera matriz de covarianza es el valor esperado de la matriz  $(\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^t$ , se obtiene un resultado muy satisfactorio.

Una vez estimados los parámetros  $\mathbf{m}$  y  $C$ , la función de densidad de probabilidad queda perfectamente especificada por la ecuación dada en (2.13) suponiendo que dicha función sigue una distribución Gaussiana.

$$p(\mathbf{x} | \mathbf{m}, C) = \frac{1}{(2\pi)^{d/2} |C|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{m})^t C^{-1} (\mathbf{x} - \mathbf{m}) \right\} \quad (2.13)$$

Según la teoría general de la probabilidad de Bayes, dado  $\mathbf{x}$  el objetivo que se plantea es asignar  $\mathbf{x}$  alguna de las clases existentes. Supongamos que las clases son  $w_1, w_2, \dots, w_c$  para cada una de ellas se puede estimar la función de densidad de probabilidad dada en (2.13) obteniendo  $(\mathbf{m}_1, C_1)$  para la clase  $w_1$ ,  $(\mathbf{m}_2, C_2)$  para la clase  $w_2$  y así sucesivamente hasta llegar a la clase  $w_c$  con  $(\mathbf{m}_c, C_c)$ .

Por tanto, dada la observación  $\mathbf{x}$  se trata de determinar la probabilidad a posteriori de que dicha observación pertenezca a la clase  $w_j$ . El proceso para llevar a cabo esta operación se puede realizar por medio de la teoría de la probabilidad de Bayes calculando la probabilidad a posteriori como sigue,

$$P(w_j | \mathbf{x}) = \frac{p(\mathbf{x} | \mathbf{m}_j, C_j) P(w_j)}{\sum_{j=1}^c p(\mathbf{x} | \mathbf{m}_j, C_j)} \quad (2.14)$$

El numerador del segundo miembro de la ecuación (2.14) incluye dos términos, a saber: 1) la función de densidad de probabilidad  $p(\mathbf{x} | \mathbf{m}_j, C_j)$  que está dada en la ecuación (2.13) y 2) la probabilidad a priori  $P(w_j)$  de que dicha muestra pertenezca a la clase  $w_j$ . Uno de los problemas que se plantean respecto del cálculo de la probabilidad a priori consiste en cómo determinar esta probabilidad antes de observar la muestra.

En cualquier caso, la decisión final sobre la asignación de  $\mathbf{x}$  a una clase dada se basa en el siguiente criterio, según Duda y col. (2001):

$$\mathbf{x} \in w_j \quad \text{si} \quad P(w_j | \mathbf{x}) > P(w_k | \mathbf{x}) \quad \forall k \neq j \quad (2.15)$$

Teniendo en cuenta la ecuación (2.14) este criterio se puede expresar como sigue ya que el denominador en (2.14) representa la densidad de probabilidad mixta, que no tiene carácter discriminante puesto que es la misma en todos los casos.

$$\mathbf{x} \in w_j \quad \text{si} \quad p(\mathbf{x} | w_j) P(w_j) > p(\mathbf{x} | w_k) P(w_k) \quad \forall k \neq j \quad (2.16)$$

Como se ha mencionado anteriormente, la única cuestión pendiente estriba en el cómputo de la probabilidad a priori. Cuando ésta no se conoce, lo que se hace es fijarla para todas las clases a un valor constante, por ejemplo a 0.5 para que no intervenga realmente en la decisión.

### 2.3.3. Algoritmo Generalizado de Lloyd

A continuación vamos a describir el algoritmo generalizado de Lloyd (AGL) propuesto inicialmente por Lloyd (1957) y generalizado más tarde para cuantización vectorial por Linde y col. (1980). Este algoritmo aplica las dos condiciones necesarias a los datos de entrenamiento para determinar cuantizadores vectoriales óptimos minimizando el riesgo empírico, el algoritmo que exponemos aquí utiliza un método de aproximación estocástica. En realidad el método propuesto a continuación es una versión modificada del AGL original y se conoce como algoritmo de aprendizaje competitivo en la literatura de redes neuronales.

1) Inicio: dados los puntos de datos  $\mathbf{x}(k)$ ,  $k = 1, 2, \dots$ , y centros de salida iniciales  $\mathbf{c}_j(0)$ ,  $j = 1, \dots, m$

2) Determinar el centro  $\mathbf{c}_j(k)$  más próximo al punto  $\mathbf{x}(k)$

$$j = \arg \min_j L(\mathbf{x}(k), \mathbf{c}_j(k)) \quad (2.17)$$

utilizando la función de error al cuadrado como función de pérdidas. Esto se simplifica para obtener la regla del vecino más próximo

$$j = \arg \min_j \|\mathbf{x}(k) - \mathbf{c}_j(k)\| \quad (2.18)$$

El hallazgo del centro más próximo se denomina competición (entre centros) en los métodos de redes neuronales.

3) Actualizar el centro de salida utilizando las ecuaciones,

$$\begin{aligned} \mathbf{c}_j(k_j+1) &= \mathbf{c}_j(k_j) - \gamma(k_j) \text{grad } L(\mathbf{x}(k), \mathbf{c}_j(k_j)) \\ k_j &= k_j + 1 \end{aligned} \quad (2.19)$$

Obsérvese que cada centro puede tener su propia razón de aprendizaje, lo que se indica con  $k_j$  en  $\gamma(k_j)$ , con  $j = 1, \dots, m$ . La razón de aprendizaje  $\gamma(k_j)$  debe cumplir las condiciones de aproximación estocástica. Cuando la función de pérdidas es el error al cuadrado, el gradiente se calcula como,

$$\frac{\partial L(\mathbf{x}, \mathbf{c}_j)}{\partial \mathbf{c}_j} = \frac{\partial}{\partial \mathbf{c}_j} \|\mathbf{x} - \mathbf{c}_j\|^2 = 2(\mathbf{x} - \mathbf{c}_j) \quad (2.20)$$

con este gradiente, los centros de salida se actualizan por

$$\begin{aligned} \mathbf{c}_j(k_j+1) &= \mathbf{c}_j(k_j) - \gamma(k_j) [\mathbf{x}(k) - \mathbf{c}_j(k_j)] \\ k_j &= k_j + 1 \end{aligned} \quad (2.21)$$

que se conoce comúnmente como regla de aprendizaje competitivo en redes neuronales.

Un problema común en el AGL es que una mala elección de los centros iniciales desemboca en soluciones óptimas localmente malas. Una aproximación consiste en aplicar el AGL varias veces comenzando con distintas condiciones iniciales y luego eligiendo la mejor solución. En redes neuronales se han propuesto también varias soluciones de entre las que destaca el “self-organizing map”. Otro método es el denominado mecanismo de consciencia (DeSieno 1988), que es realmente una modificación del procedimiento descrito por la ecuación (2.18). Cada unidad mantiene una traza del número o frecuencia de las veces que han vencido la competición en el paso 2. Sea  $frec_j(k)$  dicha frecuencia correspondiente a la unidad  $j$  en la iteración  $k$ . Entonces la regla del vecino más próximo (2.18) se modifica a

$$j = \arg \min_j \left[ \|\mathbf{x}(k) - \mathbf{c}_j(k)\| frec_j(k) \right] \quad (2.22)$$



El paso 3 no cambia. La nueva medida de distancia en (2.22) fuerza a cada unidad a vencer en promedio el mismo número de veces. En otras palabras, las vencedoras más frecuentes se sienten culpables (toman consciencia) y reducen su tendencia a vencer mediante (2.22).

## 2.4. Métodos lineales de extracción de bordes

En primer lugar se ha de definir lo que es un borde. Un borde no es más que la frontera entre dos regiones con propiedades de nivel de gris distintas. La extracción de bordes es una aplicación que ha de realizarse localmente, es decir, no sólo necesitamos el píxel donde se ha de calcular el borde, sino que necesitamos además los píxeles adyacentes.

Aprovechando que la derivada mide la variación de una magnitud, realizamos la detección de bordes a partir de un operador local de derivación. Antes de proseguir vamos a realizar algunas puntualizaciones:

- La primera derivada es distinta de 0 siempre que haya variación de nivel de gris entre píxeles adyacentes.
- En las zonas de nivel de gris constante, la derivada es nula.

La señal a la que hemos de aplicar la derivada es una imagen. Como es un objeto bidimensional lo que realmente aplicamos es lo que denominamos gradiente:

$$\Delta f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} \quad (2.23)$$

donde  $G_x$  y  $G_y$  son las derivadas parciales con respecto al eje vertical y horizontal respectivamente. La anterior expresión lleva implícita la dirección en la que queremos calcular el gradiente. Puesto que el factor que más nos interesa es la amplitud del gradiente para determinar si hay borde o no, trabajaremos mejor con la siguiente expresión:

$$G = \sqrt{(G_x)^2 + (G_y)^2} \quad (2.24)$$

Los diferentes métodos de detección de bordes que se verán implementados se derivan de distintas aproximaciones numéricas al gradiente.

## 2.5. Operadores de extracción de bordes

### 2.5.1. Roberts

Una de las opciones que se pueden implementar para detectar un borde es estudiar los niveles de gris de píxeles adyacentes y observar si hay variaciones. La forma más intuitiva de comprobar si hay variación es la diferencia. Por ello, este método implementa una máscara que, aplicada a un píxel, calcula la diferencia entre el nivel de gris de este píxel y un píxel vecino. Las máscaras usadas en el método de Roberts son:

**Gradiente Horizontal**

$$G_x = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

**Gradiente Vertical**

$$G_y = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

### 2.5.2. Prewitt

El método de Roberts nos da una primera aproximación a la hora de calcular bordes, pero es muy dependiente de la orientación del borde (hacemos la diferencia de dos píxeles que están situados sobre una dirección en el espacio). En el método de Prewitt lo que se intenta es ponderar el valor de la diferencia respecto a un número mayor de píxeles vecinos, lo que da una idea mejor de lo que hay alrededor del punto donde estamos haciendo el cálculo. Las máscaras usadas en el método de Prewitt son:

**Gradiente Horizontal**

$$G_x = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

**Gradiente Vertical**

$$G_y = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Se puede observar que ahora en ambas máscaras no se hace uso del valor central. Esto indica que para calcular un borde en un punto de la imagen, se tiene en cuenta sólo a los píxeles vecinos.

### 2.5.3. Sobel

Este método se deriva del método anterior. Si nos fijamos en alguna máscara del método de Prewitt nos damos cuenta que da la misma importancia a los píxeles que se encuentran sobre el eje vertical u horizontal que a los que se encuentran en la diagonal. El método de Sobel intenta compensar este efecto dándole una mayor importancia a los píxeles que están más cerca del centro de la máscara, es decir, a los píxeles que se encuentran sobre el eje vertical u horizontal. Las máscaras que implementa este método para aproximar el gradiente son:

**Gradiente Horizontal**

$$G_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

**Gradiente Vertical**

$$G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

## 2.6. Descriptores estadísticos de textura

Hay varios caminos para caracterizar una imagen. Se mencionarán algunas propiedades estadísticas útiles.

### 2.6.1. Media y desviación estándar

El nivel de gris medio,  $G$ , y su desviación estándar o típica,  $\sigma$ , se determinan fácilmente mediante las siguientes expresiones:

$$G = \frac{1}{F \cdot C} \sum_{x=0}^{F-1} \sum_{y=0}^{C-1} g(x, y) \quad (2.25)$$

$$\sigma = \frac{1}{F \cdot C} \sum_{x=0}^{F-1} \sum_{y=0}^{C-1} (g(x, y) - G)^2 \quad (2.26)$$

$F$  y  $C$  representan el número de filas y columnas del área considerada. La media  $g$  indica el brillo o luminosidad de una imagen mientras que la desviación estándar es la medida del contraste, o lo que es igual la variación de la información en la imagen. Un valor pequeño indica una imagen “plana” con poco contraste, es decir, una imagen con un contenido de información reducido.

### 2.6.2. Entropía

La entropía mide la aleatoriedad de un conjunto de variables aleatorias. Aplicada a una imagen digital, la entropía expresa la incertidumbre de un valor de gris. Sean  $p_i$ ,  $i = 0, 1, \dots, \text{máx}$ ; las probabilidades de los valores de grises de una imagen digital. Se define la entropía  $H$  como:

$$H = - \sum_{i=0}^{\text{máx}} p_i \cdot \log_2 p_i \quad (2.26)$$

Una imagen binaria tiene 2 niveles de gris. Suponiendo que la probabilidad de cada valor es de 0,5, la entropía será 1, lo cual tiene sentido ya que se necesita un bit por píxel para representar una imagen binaria. Para una imagen típica con 256 niveles de gris la entropía es 8, suponiendo para todos los niveles de gris la misma probabilidad de 1/256.

## 2.7. Identificación de líneas rectas: transformada de Hough

Se trata de un método propuesto por Hough (1962) para encontrar la ecuación de una línea que pase por un conjunto de  $n$  puntos en el plano  $xy$ . Una línea que une una secuencia de puntos o píxeles puede expresarse de la forma,

$$y = ax + b \quad (2.27)$$

el objetivo es evaluar los parámetros  $a$  y  $b$ .

Considérese un punto  $(x_i, y_i)$ , existe un número infinito de líneas que pasan por dicho punto y que satisfacen todas la ecuación  $y_i = ax_i + b$  para valores variables de  $a$  y  $b$ . Si escribimos la ecuación como  $b = -x_i a + y_i$  y consideramos el plano  $ab$  también llamado espacio de parámetros, tendríamos la ecuación de una única línea para un valor fijo de  $(x_i, y_i)$ . Un segundo punto  $(x_j, y_j)$  tendrá también una línea en el espacio de parámetros asociada y esta línea interceptará con la línea asociada a  $(x_i, y_i)$  en  $(a', b')$ , donde  $a'$  es la pendiente y  $b'$  la altura a la que la recta que contiene a los dos puntos  $(x_i, y_i)$  corta al eje  $xy$ . De hecho todos los puntos contenidos en esta línea tendrán rectas asociadas en el espacio de parámetros que se cortarán en  $(a', b')$ . El ejemplo de la tabla 2.1 y la figura 2.3 ilustra estos conceptos

$x$	$y$	espacio $xy$	espacio $ab$
1	0	$0 = a + b$	$b = -a$
3	2	$2 = 3a + b$	$b = -3a + 2$
4	3	$3 = 4a + b$	$b = -4a + 3$
3	4	$4 = 3a + b$	$b = -3a + 4$

Tabla 2.1 Valores en los espacios  $xy$  y  $ab$

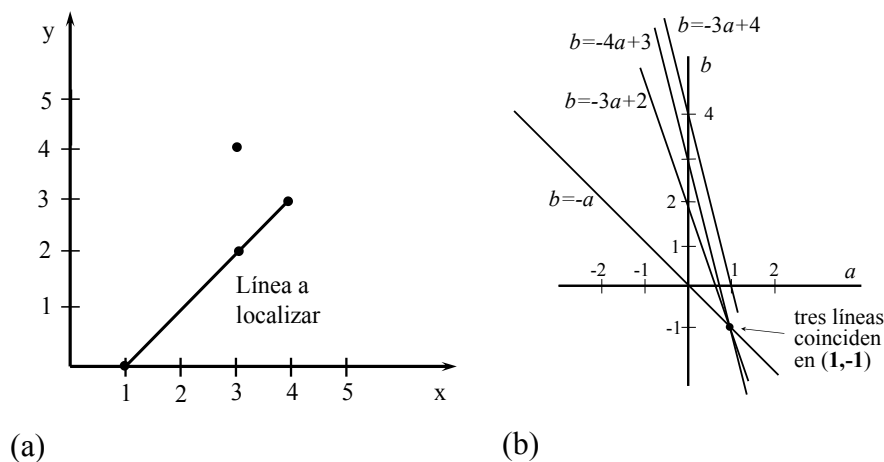


Figura 2.3 Ilustración de la transformada de Hough. (a) Recta original en el espacio  $xy$  con tres puntos de la misma. (b) Acumulador en el espacio de parámetros  $ab$ , los tres puntos de la recta en  $xy$  se transforman en tres líneas que se cortan en un punto  $(1, -1)$  en el espacio de parámetros  $ab$

Pasemos a usar la transformada de Hough para nuestro problema. Subdividimos el espacio de parámetros en las llamadas celdas de acumulación, como se muestra en la figura 8.6, donde  $(a_{max}, a_{min})$  y  $(b_{max}, b_{min})$  son los intervalos dentro de los cuales estarán la pendiente y

la ordenada en el origen. La celda de acumulación  $A(i,j)$  se corresponde con un cuadrado asociado a las coordenadas  $(a_j, b_j)$ . Estas celdas se inicializan a cero. Para cada punto  $(x_k, y_k)$  del plano de la imagen, tomamos como  $a$  a cada una de las subdivisiones del eje  $a$  y se calcula el  $b$  correspondiente usando la ecuación  $b = -x_k a + y_k$ . Los valores de  $b$  obtenidos se redondean al valor más cercano del eje  $b$ . Si al tomar  $p$  obtenemos  $q$ , tendremos  $A(p,q) = A(p,q) + 1$ . Cuando se acabe con este proceso, un valor  $M$  en la celda  $A(i,j)$  se corresponde con  $M$  puntos del plano  $xy$ , que se encuentran en la recta  $y = a_i x + b_j$ . Como se menciona en Fu y col. (1988), la exactitud de la colinealidad de estos puntos viene dada por el número de subdivisiones establecidas en el plano  $ab$ .

Si se divide el eje  $a$  en  $K$  partes, para cada punto  $(x_k, y_k)$  se obtienen  $K$  valores de  $b$  correspondientes a los  $K$  posibles valores de  $a$ . Como existen  $n$  puntos de imagen, se necesitan  $nK$  cálculos.

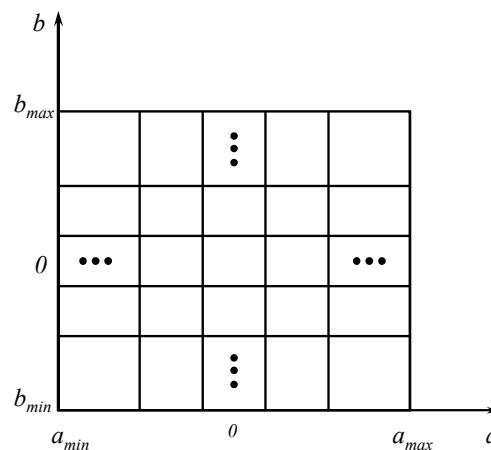


Figura 2.5 División del plano de parámetros para su uso en la transformada de Hough

Un problema importante surge a la hora de fijar los límites  $a_{max}$ ,  $a_{min}$  para recorrer el rango definido por estos valores, ya que a priori no son conocidos. Otro problema que aparece al utilizar la ecuación  $y=ax+b$  para representar una recta es que tanto la pendiente como la ordenada en el origen se acercan a infinito cuando la recta se aproxima a posiciones verticales. Una forma de evitar esto es utilizar la representación normal de la recta, dada por  $x \cos \theta + y \sin \theta = \rho$ , ahora el espacio de parámetros viene determinado por  $\theta$  y  $\rho$ . Esto posibilita barrer todo el rango de valores para el ángulo  $\theta$  desde  $-180^\circ$  a  $+180^\circ$ .

La complejidad de la transformada de Hough depende en gran medida, del número de coordenadas y coeficientes de la función a la que se quiera aplicar.

La transformada de Hough puede usarse para reconocer una forma arbitraria, que es conocida como la transformada de Hough generalizada. Este método es esencialmente una forma de descripción de bordes. El método consiste en elegir un punto dentro del objeto, por ejemplo el centroide. Luego, siguiendo el borde del objeto, y para cada punto de dicho borde o frontera se obtiene la distancia al punto central  $r$  con ángulo  $\phi$ . El conjunto de datos  $r, \phi$  se denomina R-tabla. La transformada de Hough ha sido probada en presencia de ruido o cuando hay partes ocultas o no presentes en el objeto.





## 3. Diseño de algoritmos

### 3.1. Introducción

Antes de comenzar con el proceso de diseño de un algoritmo conviene destacar las fases o mecanismos involucrados en cualquier proceso en el cual esté involucrada la visión por computador. La figura 3.1 muestra un esquema general donde se detalla esto.

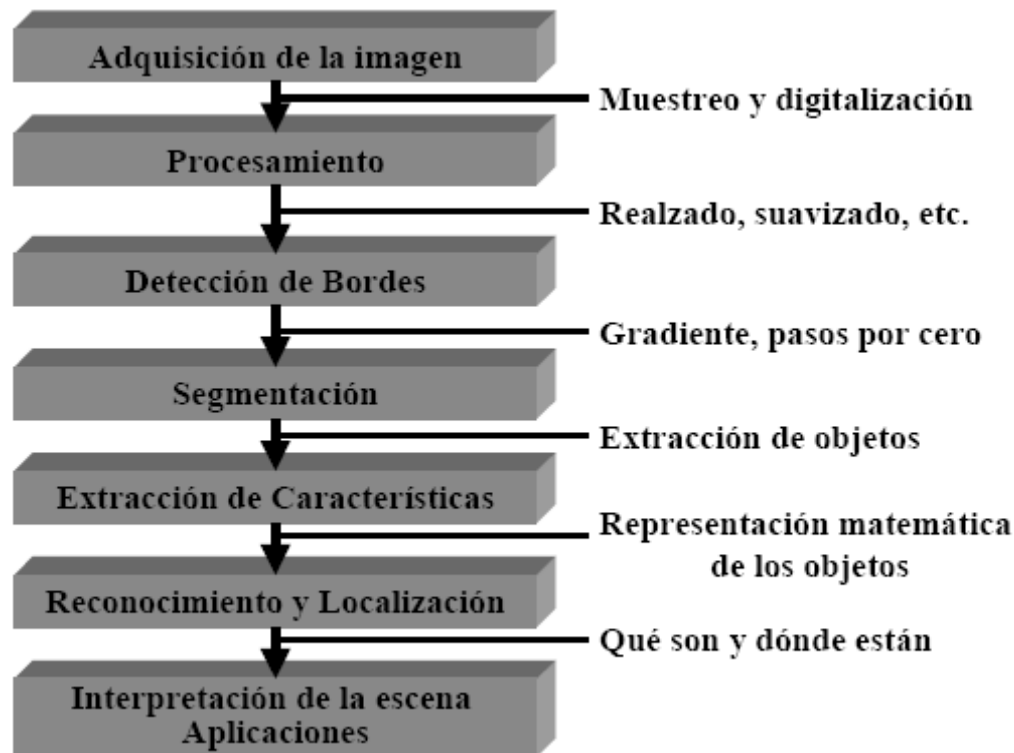


Figura 3.1 Esquema de un algoritmo de visión por computador

Este proceso es la base sobre la cual se construirán la práctica totalidad de los algoritmos orientados a resolver problemas de visión por computador. Es posible, que algunos casos el algoritmo diseñado no tenga que aplicar ciertas fases del proceso. Por ejemplo, puede ocurrir que las imágenes de las que disponemos sean de gran calidad y no necesiten de ningún método de pre-procesamiento o que simplemente no me interese saber de que tipo son los objetos que componen la escena.

A continuación, se detallan brevemente los pasos a seguir durante cada una de las fases y su relación con el método que se presenta en este trabajo:

***Captura de imágenes:*** se obtienen las imágenes a partir de una fuente de imágenes existentes o bien a partir de la captura directa de la escena mediante los correspondientes dispositivos de captura. En nuestro caso son proporcionadas en formato JPEG y capturadas mediante el dispositivo estereoscópico con número de patente nacional MU-200501738. Se trata de un sistema que proporciona imágenes en perspectiva circular conocida como dispositivos de ojo de pez.

***Procesamiento:*** En esta fase se tratan las operaciones y transformaciones que se aplican sobre las imágenes digitales en una etapa de procesamiento previa a las de segmentación y reconocimiento. Su objeto es mejorar o destacar algún elemento de las imágenes, de manera que sea posible realizar las etapas posteriores de segmentación y clasificación. En nuestro caso, esta etapa no ha sido necesaria.

***Detección de bordes / Segmentación:*** La segmentación es un proceso que consiste en dividir una imagen digital en regiones homogéneas con respecto a una o más características (como por ejemplo el brillo o el color) con el fin de facilitar un posterior análisis o reconocimiento automático. En nuestro caso dentro de esta etapa se encuentra la extracción de bordes y el procesado de la imagen en color.

***Extracción de características:*** Una vez segmentada una imagen es importante considerar la forma de describir los objetos localizados. En esta fase se intentará que usando descripciones apropiadas pueda conseguirse una representación única e independiente de la posición, la orientación y el tamaño del objeto descrito. En el método propuesto en este trabajo a esta parte pertenecen los denominados descriptores de textura.

***Reconocimiento:*** con esta información se pueden aplicar los métodos de aprendizaje que se deseen, para tratar de determinar que es el objeto representado. También se pueden aplicar modelos matemáticos para tratar de identificar patrones o figuras. Es la fase donde se encuadran los métodos de clasificación.

*Interpretación de la escena:* se corresponde con la etapa final de obtención del volumen de madera, quedando fuera del objetivo del presente trabajo.

## **3.2. Identificación del problema**

La motivación para el diseño de un algoritmo que permita calcular alturas y volúmenes de árboles sobre sistemas estereoscópicos viene de la necesidad de controlar el crecimiento de ciertas especies y determinar el volumen de madera existente en una determinada extensión o área. Como se ha mencionado previamente, esta tarea actualmente se realiza de forma manual requiriendo un elevado tiempo de proceso. El objetivo final es su automatización, siendo este trabajo uno de los pioneros en este sentido. Por tanto, se plantea la necesidad de un método que consiga resultados precisos de manera automática.

Más concretamente el objetivo consiste en aislar dentro de las imágenes exactamente los troncos de los árboles para poder extraer las propiedades anteriores con las que establecer la correspondencia en un sistema estéreo.

El aislamiento consistirá en identificar para su discriminación todo lo que no sea tronco del árbol, es decir: suelo, cielo y hojas principalmente.

Se dispone de varios tipos de imágenes, que se describen a continuación, donde se analiza la problemática relacionada con las mismas.

Los objetivos por tanto son bastante claros:

1. Determinar el tamaño del tronco: área, altura.
2. Establecer la correspondencia entre los troncos de los árboles de las dos imágenes del par estereoscópico para determinar su distancia al sistema estéreo.

Este trabajo se centra exclusivamente en el primero combinando las técnicas anteriormente descritas de la manera más convenientemente posible.

El segundo objetivo se pospone para futuras investigaciones.

En lo que resta del capítulo se realiza un análisis de la problemática subyacente en las diversas imágenes disponibles. Seguidamente se analizan los métodos de procesamiento

con el fin de conseguir el objetivo número uno. Con tal propósito, en primer lugar, se diseña un procedimiento o método basado en clasificación; en segundo lugar se describe un proceso basado en la descripción de texturas. Finalmente, en tercer lugar la propuesta se realiza en base a la extracción de bordes.

Conviene reseñar que a pesar de que en las imágenes aparece una alta densidad de árboles, principalmente en la lejanía, el objetivo del análisis se centra en las zonas próximas al lugar donde se ubica el sistema de estéreo-visión. La distancia útil o eficaz del sistema resulta ser aproximadamente de unos 25 metros de diámetro. El dispositivo se desplaza convenientemente por la superficie del bosque teniendo como referencia dicha distancia hasta cubrir la superficie total a analizar. Una consecuencia importante derivada de esta situación estriba en el hecho de que solamente es necesario identificar las texturas de los troncos de árbol más próximos y por tanto más significativos.

### **3.3. Análisis de las imágenes: problemática**

A la vista de las imágenes analizadas se identifica la problemática que a continuación se describe:

#### **1. Imágenes captadas con escasa iluminación**

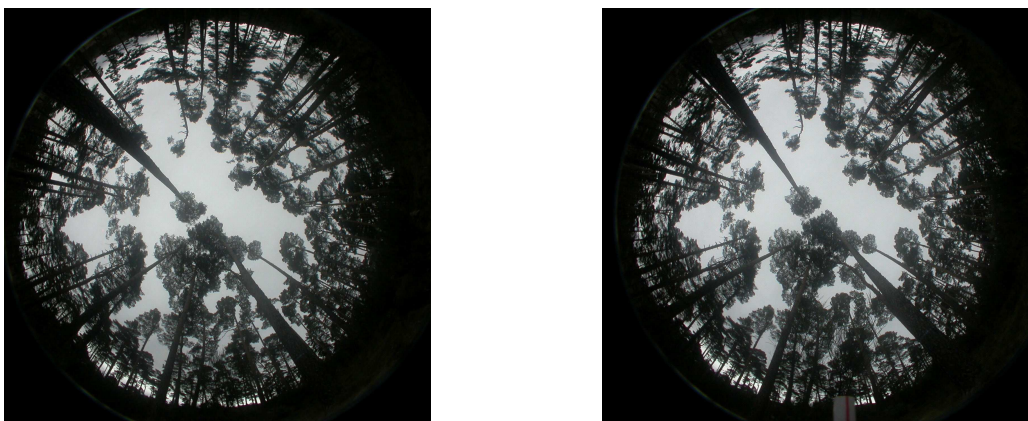


Figura 3.2 Par de imágenes estereoscópicas

**Color:**

En este caso los troncos de los árboles presentan desde la base al menos tres tonalidades de color comenzando desde las negras y evolucionando hacia copas de los árboles donde alcanza diferentes niveles o tonalidades (que resultan ser casi grises). Esto implica que un mismo tronco de árbol no posee la misma textura de color debido a los diferentes niveles de iluminación.

Las tonalidades de los troncos son similares a las del resto de texturas existentes según la situación de la vegetación en la imagen.

Conclusión: no se pueden identificar los troncos sólo por la textura de color debido a su similitud con la maleza, no así con las texturas del cielo que sí admiten una fácil discriminación

### **Descriptores de textura estadísticos:**

En este tipo de imágenes es posible distinguir a través de los momentos estadísticos (varianza) diferentes texturas, diferenciándose las texturas de las ramas de las de los troncos por presentar en este último caso una cierta variabilidad, mientras que en el primero se manifiestan de forma homogénea.

## **2. Imágenes con variedad de color**

En este tipo de imágenes resulta ya más fácil separar la parte predominantemente verde de la que no lo es. Esta última serían los troncos. Si bien, éstos presentan una problemática que es su distinta tonalidad, en efecto se pueden distinguir varias tonalidades:

Oscuras: quedan en la parte de la sombra

Claros: por recibir la iluminación solar directamente, en este caso llega a tener el mismo valor que parte del cielo, con lo que la problemática respecto de su posible clasificación se incrementa considerablemente.

Estas dos apreciaciones se manifiestan en la imagen de la izquierda y no así en la imagen de la derecha donde dado el espesor de vegetación no hay sombras y sólo cabe distinguir entre verde, azul, blanco y oscuro. Siendo la parte oscura la correspondiente a los troncos.

Se aprecian también diferentes texturas. Distinguiéndose entre las texturas en las zonas donde hay hojas y las texturas de los troncos. Las primeras presentan alta variabilidad entremezclándose verde con azul, mientras que las texturas que representan los troncos no tienen este comportamiento.



Figura 3.3 Imágenes estéreo

### 3.4. Diseño del método mediante técnicas de clasificación

El método consta de las siguientes fases:

- 1) Realizar un proceso de entrenamiento previo de forma que se extraen muestras de las texturas a identificar. Al nivel de la investigación actual, este proceso de selección se realiza de forma manual por lo que en el futuro debe procederse a su automatización.
- 2) Tras el proceso de aprendizaje se realiza la clasificación.
- 3) Calcular la desviación estándar las texturas que representan las hojas.
- 4) Eliminar las texturas de las hojas de las clases de los troncos.
- 5) Eliminar regiones con áreas pequeñas y excentricidades bajas.
- 6) Dibujar una circunferencia de radio ligeramente inferior al de la imagen.

- 7) Extraer la región que posea el área mayor y eliminar la circunferencia.

Se muestra a continuación mediante un ejemplo la aplicación los pasos que conforman el proceso previo a partir de la imagen original mostrada en la figura 3.4. Se selecciona un conjunto de puntos, cada uno de los cuales está caracterizado por las componentes de color R, G y B de suerte que el vector que lo describe será  $x = (R,G,B)$ . Estos vectores  $x$  son los patrones utilizados en los métodos de clasificación descritos en el capítulo 2. Este conjunto de puntos de entrenamiento se usará como entrada del sistema que usará el método general de Lloyd descrito en la sección 2.



Figura 3.4 Imagen original

Con este conjunto de muestras de entrenamiento (píxeles) se obtienen los centros de los clústeres mediante las técnicas de clasificación descritas en el capítulo 2. Así, mediante el método de agrupamiento borroso se obtienen los centros de las clases  $v_j$  según la ecuación (2.6). Mediante el clasificador Bayesiano los centros  $m_j$  sí como las matrices de covarianza según la ecuación (2.12). Finalmente, mediante el algoritmo de Lloyd se obtienen los centros  $c_j$  de acuerdo con la ecuación (2.21). Los parámetros así aprendidos se almacenan en la denominada base de conocimiento. En este momento finalizaría el proceso de entrenamiento. En la figura 3.5 se representa una serie de muestras, junto con sus respectivos centros, para tres clústeres de la imagen de la figura 3.4 correspondientes a tres tipos de texturas, a saber: cielo, troncos y hojas-hierbas.

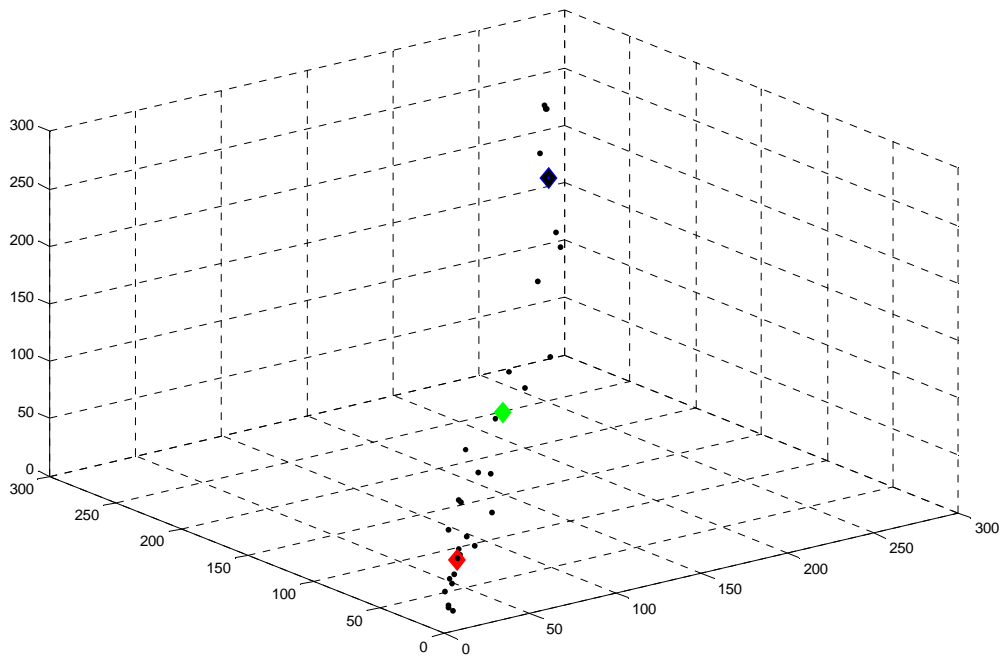


Figura 3.5 Creación de un cluster

El siguiente paso consiste en realizar una etapa de clasificación para toda la imagen según se describe a continuación:

Ante la llegada de cada nueva muestra, se trata de decidir a qué clase pertenece. Para ello se recuperan de la base de conocimiento los datos almacenados previamente durante la fase de entrenamiento.

Dada una nueva muestra  $\mathbf{x}$  para ser clasificada, se trata de decidir a qué clase de las tres mencionadas pertenece. Cada método asigna a  $\mathbf{x}$  una clase, al disponer de tres métodos, la decisión final se toma por mayoría. Se trata de una técnica clásica conocida como *votación mayoritaria* (*majority voting*) (Kittler, 1998; Kuncheva, 2004).

**Agrupamiento borroso:** Con los centros  $\mathbf{v}_j$  proporcionados, calculamos los grados de pertenencia de  $\mathbf{x}$  a cada clase, el mayor grado nos determina la clase a la que pertenece.

**Bayes:** con los centros  $\mathbf{m}_j$  y las matrices de covarianza calculadas se obtiene la probabilidad de pertenencia de  $\mathbf{x}$  a cada una de las clases según la ecuación (2.13),



decidiendo que  $x$  pertenece a la clase que proporcione la mayor probabilidad según la ecuación (2.16).

**Algoritmo de Lloyd:** mediante los centros  $c_j$  determinamos a través de la ecuación (2.22), la pertenencia de  $x$  a los clústeres.

Por tanto, cada método clasifica  $x$  como perteneciente a una clase, al disponer de tres métodos la decisión final se toma por mayoría como se ha mencionado previamente.

En la imagen 3.6 se pueden observar los resultados obtenidos mediante el procedimiento descrito anteriormente.

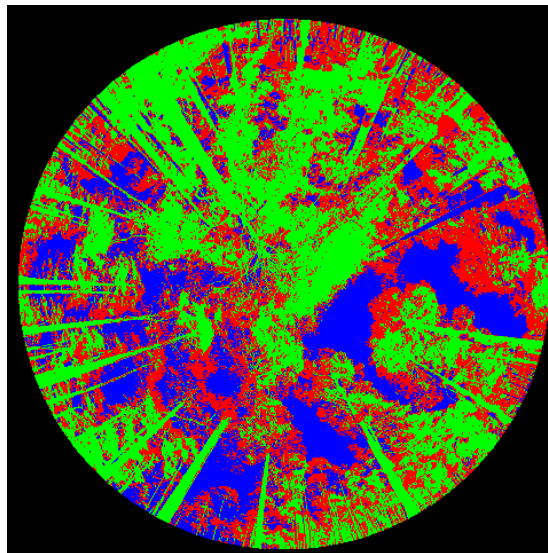


Figura 3.6 Resultados de la clasificación: troncos en verde, cielo en azul y hojas en rojo

Aunque los resultados obtenidos son aceptables, no son suficientes por lo que recurriremos a la extracción de las texturas que representan las hojas en la imagen. Para ello, calculamos la desviación estándar para cada píxel teniendo en cuenta sus ocho vecinos. Esto se lleva a cabo a través de la ecuación (2.26) con  $F = C = 3$ , que son los valores que definen dicha vecindad. Las hojas se caracterizan por ser zonas de “alta textura”. Los resultados se pueden ver en la imagen 3.7.

Con las clases de los troncos y las texturas obtenidas se procede a limpiar las texturas para aislar los troncos con la siguiente regla empírica obtenida a partir de la observación de las imágenes:

Si es una textura obtenida por mediante la desviación estándar y no ha sido clasificado como perteneciente a una de las clases que describen los troncos entonces se pone a cero, sino a blanco (tratando de dejar sólo los troncos). De nuevo se pueden ver los resultados de la técnica en la figura 3.8.

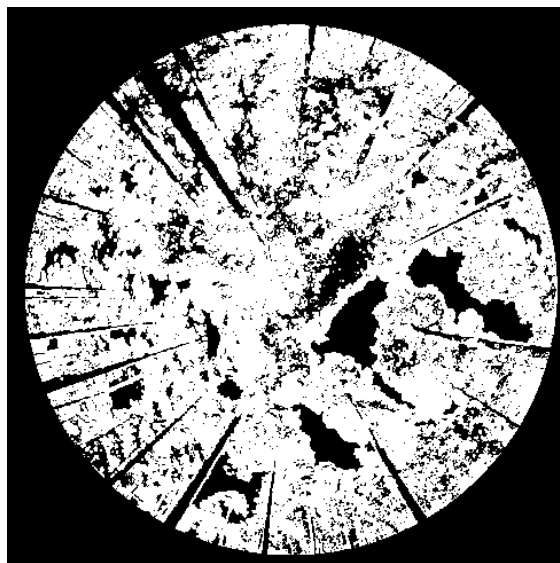


Figura 3.7 Texturas: desviación estándar

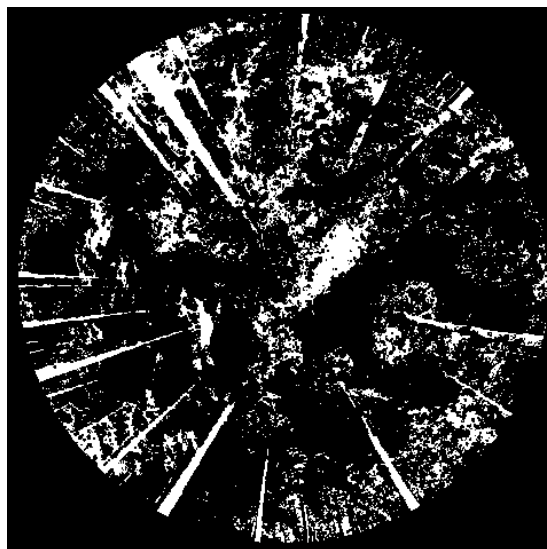


Figura 3.8 Troncos: sin hojas

Aún queda una gran cantidad de ruido en la imagen, pero si lo observamos bien vemos que está compuesto por muchos grupos de puntos que tienen áreas muy pequeñas y formas muy cercanas a elipses o círculos. Esto nos da la idea de eliminar todas las regiones conexas con un área pequeña y una excentricidad baja, obteniendo los resultados que se muestran en la imagen de la figura 3.9.

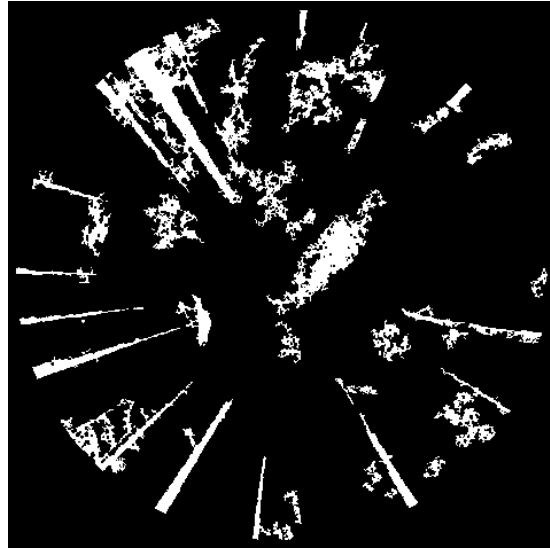


Figura 3.9 Troncos: texturas limpiadas

El proceso aunque efectivo no es suficiente, ya que existen regiones en las que el área es suficiente como para representar un árbol e incluso poseen formas similares, razón por la cual aún quedan sin eliminar un pequeño número de las mismas.

La idea consiste en dibujar una circunferencia de un radio algo menor al de la imagen. Al realizar esta operación logramos que una misma región contenga todos los troncos de los árboles y donde las regiones con hojas se han eliminado. Esto se basa en el principio de que normalmente las hojas de los árboles crecen en las zonas más altas de los mismos, por lo que cortar los troncos por su base nos da un método bastante bueno para distinguir lo que es un árbol de algo que no lo es.

En las figuras 3.10 y 3.11 se muestran los resultados obtenidos finalmente.

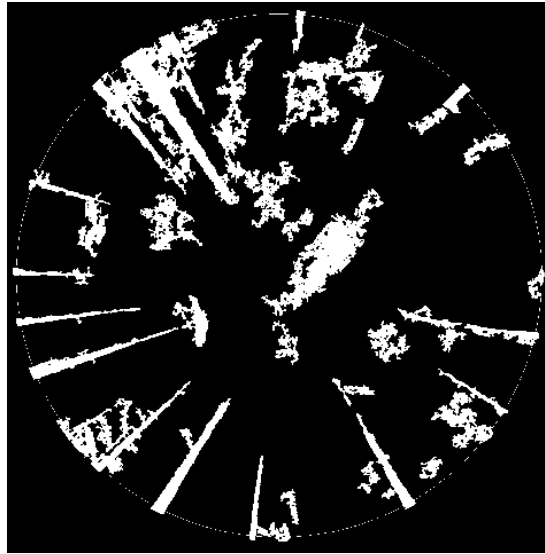


Figura 3.10 Troncos conectados por una circunferencia

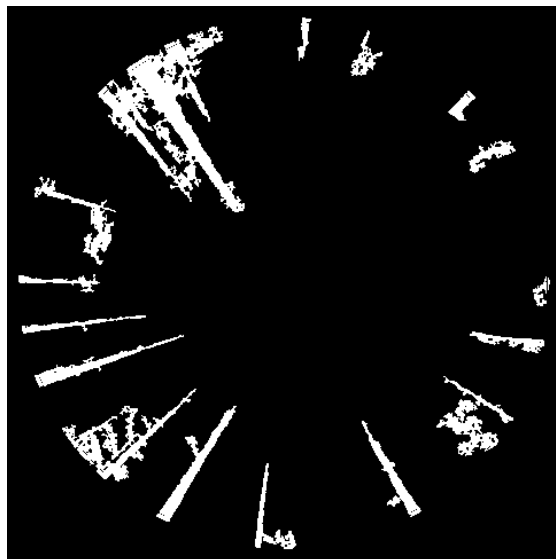


Figura 3.11 Resultado final

### 3.5. Diseño del método usando un filtro de entropía

El método se divide en las siguientes fases:

- 1) Igualar los niveles de luminosidad de la imagen (OPCIONAL)
- 2) Aplicar filtro basado en la entropía a cada píxel de la imagen.
- 3) Normalizar los valores obtenidos en el paso anterior.

- 4) Elegir un valor de corte (p.e. 0,6). Solo guardamos los valores inferiores a este valor.
- 5) Extraer la región que posea el área más grande. Esta técnica garantiza que el suelo siempre (representado como una circunferencia) contendrá todos los troncos.
- 6) Post-procesamiento extra si fuera requerido. (OPCIONAL)

La figura 3.12 muestra la imagen original que constituye el punto de partida para este método.

El primer paso es opcional y depende exclusivamente del tipo de imágenes que se están utilizando. Se ha constatado que imágenes con diferencias de iluminación demasiado grandes entre varias zonas no producen buenos resultados con este método, esto se debe, a que debemos de elegir un valor de corte para el paso 4, pero este valor no será bueno a menos que todas las regiones de la imagen tengan más o menos la misma iluminación. Para solventar este problema se recurre a dos métodos:

- Igualar los niveles de iluminación dentro de la imagen.
- Aplicar este método sobre regiones de la imagen y no sobre la totalidad, esto nos permite adaptar el valor de corte a la cantidad de iluminación presente en la región.



Figura 3.12 Imagen Original

El segundo paso esta basado en el concepto de entropía definido en la sección 2, ecuación (2.26). La idea es generar un filtro que devuelva una matriz, donde cada píxel contiene la

entropía de los 3x3 vecinos alrededor del correspondiente píxel en la imagen. Este filtro esta pensado para aplicarse exclusivamente a imágenes con 256 niveles de gris. Si la imagen fuera a color sería necesario transformarla a escalas grises. La matriz devuelta por el filtro de entropía deberá tener el mismo tamaño que la imagen original. Para los pixeles situados en los bordes se usará simetría. Tras la aplicación de este paso se obtiene una matriz de entropía que denominamos J. En la figura 3.13 se muestra el resultado de este proceso.

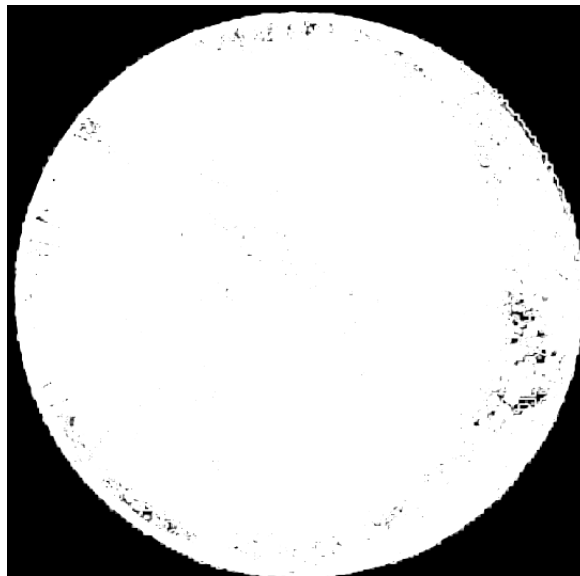


Figura 3.13 Imagen después de aplicar el filtro de entropía

En este paso también es posible usar en vez de la entropía otras medidas como la desviación estándar, ecuación (2.25).

El siguiente paso consiste en la normalización de la matriz J, obtenida mediante el filtro de entropía. Para ello se aplica una función para normalizar los valores de la matriz de intensidad J, a valores entre [0,1]. Obteniendo así una matriz resultante que denominamos E. En la figura 3.14 se muestra el resultado tras la normalización de entropía.

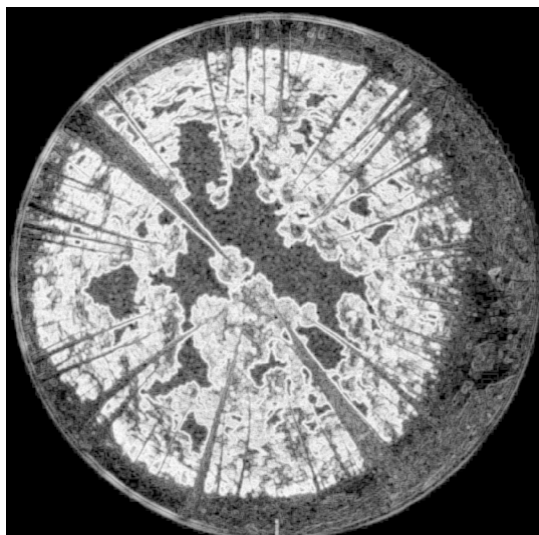


Figura 3.14 Imagen después de normalizar la entropía

El siguiente paso a realizar consiste en binarizar la imagen. Para ello se debe elegir un valor de corte apropiado. El elegir un valor de corte apropiado es crucial. La idea de este paso es muy simple, y se resume en una función que se comporte de la siguiente manera:

- Si el valor está entre (0, valor de corte) la función devolverá 1.
- En caso contrario devolverá 0.

La figura 3.15 muestra el resultado de la binarización sobre la imagen de la figura 3.14 tras el proceso previo. Actualmente, el umbral se fija manualmente estando previsto en el futuro la obtención de un umbral óptimo por ejemplo mediante el método de Otsu (1979).



Figura 3.15 Imagen binarizada

Llegados a este punto es necesario eliminar una gran cantidad de “ruido” producido fundamentalmente por fragmentos de hojas. Como el suelo siempre formará una circunferencia nos aseguramos que todos los troncos estarán conectados, esto se traduce, en que el área más grande existente en la imagen será inequívocamente la circunferencia con todos los troncos.

Gracias a este método se consigue limpiar la imagen lo suficiente como para seguir trabajando con ella con ciertas garantías de éxito. Posteriormente, si el resultado aún no fuera lo suficientemente bueno se puede aplicar un nuevo filtro para eliminar regiones con áreas demasiado pequeñas como para ser tratables. En la figura 3.16 se muestra la imagen resultante del proceso global.



Figura 3.16 Resultado del proceso

### 3.6. Diseño de un método mediante extracción de bordes

El método se divide en las siguientes fases:

- 1) Convertir la imagen de entrada a escala de grises.
- 2) Aplicar un filtro de mediana para reducir el ruido de la imagen.
- 3) Aplicar el algoritmo de Sobel para la extracción de bordes.
- 4) Reducir el número de componentes no conexas.



- 5) Extraer la imagen complementaria.
- 6) Aplicar una operación XOR a nivel de bit entre la imagen del paso anterior y la extracción de otra imagen que solo contenga el cielo.

Partiendo de la imagen original de la figura 3.17, lo primero que deberíamos hacer es transformarla del dominio RGB al de escalas grises. Como la información del color no será usada durante este proceso, es mejor hacer la transformación de la imagen a escalas grises, ya que nos permitirá trabajar de manera más cómoda.



Figura 3.17 Imagen original

Una vez transformada la imagen a escalas grises, usamos un filtrado de la mediana. Dicho filtrado es una operación no lineal que se usa normalmente para reducir el ruido de “sal y pimienta” en las imágenes. También sería posible usar convolución, sin embargo, el filtrado de la mediana tiene un rendimiento mejor cuando tratamos de disminuir el ruido, preservando los bordes.

En las distintas pruebas realizadas parece que con aplicar el filtro de la mediana sobre una matriz de 3x3 vecinos es suficiente, y con vecindades mayores no se obtienen mejoras significativas en procesamiento posterior.

Una vez hemos reducido el posible ruido existente en la imagen se aplica el operador de Sobel que se describió en el capítulo 2.4.3. En la figura 3.18 se puede ver el resultado de aplicar dicho operador.

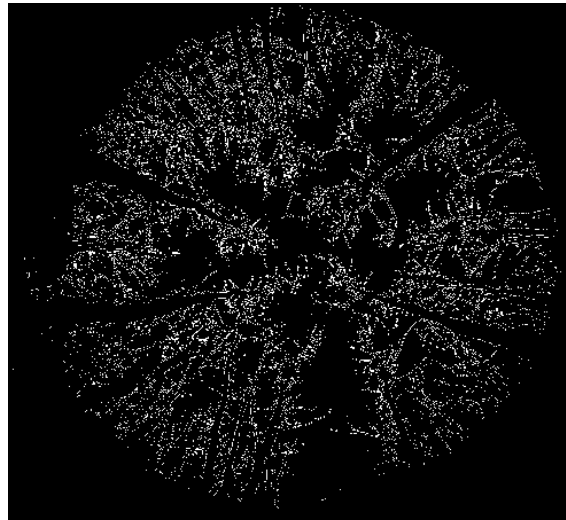


Figura 3.18 Imagen después de aplicar Sobel

Como se aprecia en la imagen de la figura 3.18, el resultado no es lo suficientemente bueno como para proseguir con los procesos siguientes, razón por la que se usa convolución sobre esta imagen. La idea es aplicar convolución sobre la imagen junto con la siguiente máscara:

$$M = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

En la figura 3.19, se observa cómo este proceso mejora significativamente la calidad del resultado y aumenta la precisión, haciendo que en esta ocasión se encuentren los bordes más definidos. Una vez disponemos de una imagen de la suficiente calidad, necesitamos obtener la imagen complementaria de la misma, pero sólo del círculo que contiene

información sobre los árboles, y no de toda la imagen. El resultado final de este paso se muestra en la figura 3.20.

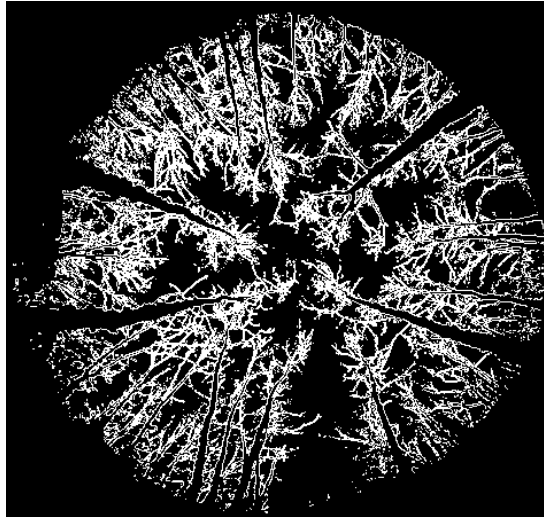


Figura 3.19 Imagen después de aplicar convolución en 2-D

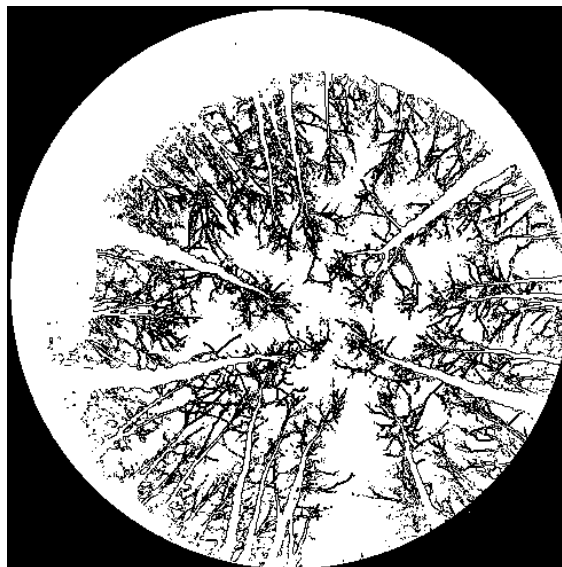


Figura 3.20 Imagen complementaria a la figura 3.19

Puesto que el cielo por norma general, es la textura más fácil de diferenciar la idea para el siguiente paso es muy sencilla. Lo primero que se hace es extraer de la imagen original la componente que representa el cielo (mayor intensidad en la imagen). Una vez disponemos de la región que representa el cielo aplicaremos la operación XOR a nivel de bit entre el resultado mostrado en la figura 3.20 y la mencionada imagen.. El resultado final se muestra

en la imagen de la figura 3.21 donde puede observarse que los resultados obtenidos para esta imagen mediante las técnicas descritas son cualitativamente hablando satisfactorias.

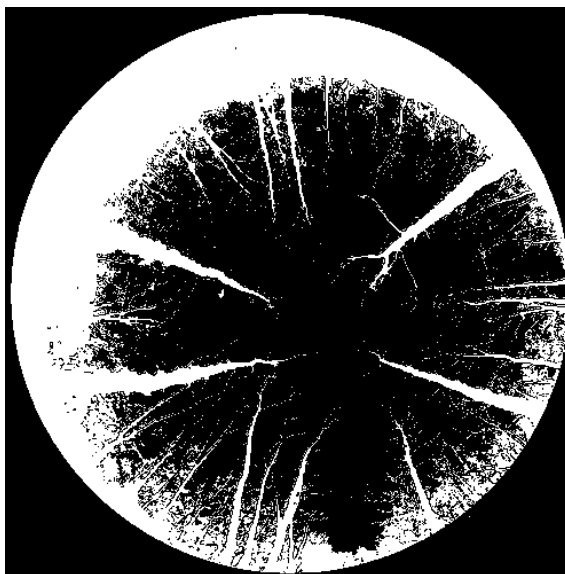


Figura 3.21 Resultado final del proceso

## 4. Análisis de resultados

### 4.1. Objetivos del análisis

En este punto se trata de verificar y validar las estrategias propuestas. Puesto que, como ya se ha mencionado, no existe investigación específica relevante en este área, no resulta posible comparar los métodos propuestos con otras estrategias. Por otra parte y dado el grado de avance del proyecto, tampoco es posible obtener resultados cuantitativos verificables por los técnicos forestales del INIA y la ETSIM debido a que falta completar el proceso mediante el análisis estereoscópico, que se plantea para futuros trabajos de investigación. Por consiguiente el principal objetivo será analizar las ventajas e inconvenientes de cada uno de los métodos propuestos.

En cuanto a la precisión de los resultados, es difícil establecer cual es el mejor en determinadas ocasiones, ya que el único método conocido de validación es el ojo humano que es una medida subjetiva.

### 4.2. Descripción de las imágenes utilizadas

Se dispone de un pequeño conjunto de imágenes adquiridas por la Escuela de Ingenieros de Motes. Se trata de imágenes digitales estéreo comprimidas en formato JPG y con una dimensión de 2272x1704 píxeles. Los sistemas estéreo producen un par de imágenes con un pequeño desplazamiento entre una y otra, este sistema es parecido al que usamos los seres humanos en el que captamos dos imágenes con un pequeño desplazamiento para poder calcular las distancias en la escena 3D.

Seleccionaremos 4 imágenes del conjunto inicial para realizar los test con los distintos algoritmos de procesamiento. A este conjunto lo llamaremos conjunto de pruebas.

A continuación, se muestran en las figuras 4.1 y 4.2 un par de imágenes del conjunto de pruebas con sus correspondientes pares estéreo. Estas imágenes son bastante complejas por la gran cantidad de elementos presentes en las mismas y que de manera muy frecuente se cruzan unas con otros. Atendiendo al problema, podríamos clasificar las texturas existentes como:

- Troncos de los árboles. Habitualmente en marrón
- Hojas de los árboles. Normalmente en verde
- Cielo. Representado por un color que va desde el negro al azul claro
- Suelo. Puede ser de cualquier color, pero normalmente será marrón o verde



(a)



(b)

Figura 4.1 (a) imagen izquierda; (b) imagen derecha



(a)



(b)

Figura 4.2 (a) imagen izquierda; (b) imagen derecha

## 4.3. Análisis de resultados

### 4.3.1. Criterio de validación

Debido a la mencionada inexistencia de métodos conocidos para la comparación se hace necesario establecer para cada una de los algoritmos una comprobación. El objetivo por tanto, se centra ahora en establecer un procedimiento para poder llevar a cabo el proceso de validación.

Queda también, por lo mencionado previamente, fuera del alcance de este trabajo el estudio o creación de un test que obtenga una medida cuantitativa de los algoritmos. De esta manera, trataremos de centrar nuestro análisis en las diferencias existentes en cada uno de los algoritmos, explicando sus ventajas y sus inconvenientes. La calidad de los resultados obtenidos se evalúa de forma cualitativa atendiendo a criterios subjetivos de los técnicos forestales..

A continuación, analizamos los tres métodos describiendo sus principales ventajas e inconvenientes.

#### *Algoritmo de Clasificación*

Las principales ventajas de este algoritmo son:

- Capacidad de aprendizaje.
- Aplicable sin preprocesamiento de las imágenes.
- Obtiene buenos resultados en todas las imágenes.

Los principales inconvenientes son:

- Tiempo de aplicación del algoritmo muy alto. Pueden llegar a ser minutos.
- Es un método de aprendizaje supervisado. Es decir, será necesario seleccionar manualmente un conjunto inicial de puntos de la imagen. En el futuro se plantea su evolución hacia la no supervisión.

Este método a priori resulta el más prometedor, ya que consigue resultados aceptables y se puede usar siempre con todas las imágenes sin grandes modificaciones. Su principal

problema es que es un algoritmo muy costoso computacionalmente hablando, particularmente si se compara con los otros algoritmos. Su otro problema puede llegar a desaparecer una vez entrenemos al sistema con suficientes imágenes, sin embargo es cierto que las especies de árboles varían mucho entre ellas, por lo que la propuesta consiste en realizar el aprendizaje por especies arbóreas para maximizar los resultados.

Cuando se aplica en imágenes de color, este método mejora considerablemente sus resultados.

### *Algoritmo usando la entropía para texturas*

Las principales ventajas de este algoritmo son:

- Gran rapidez.
- Simplicidad.
- Funciona muy bien en imágenes oscuras, obteniéndose resultados con un alto grado de satisfacción aparente.

Los principales inconvenientes son:

- Mal funcionamiento si no se trata de igualar el nivel de iluminación en todas las zonas de la imagen.
- En general, extrae de manera muy precisa los troncos de los árboles pero también superficies como el suelo que no aportan nada a la solución del problema.
- Elección de un punto de corte apropiado para cada imagen.

Este método destaca por su simplicidad y rapidez pero presenta ciertas carencias con imágenes que posean distintas intensidades luminosas. Este problema se puede solucionar intentando igualar la intensidad de la imagen en todos sus puntos o bien procesando la imagen por sectores aplicando puntos de corte distintos para cada zona.

La elección de un punto de corte adecuado hará que el algoritmo maximice los resultados obtenidos. Aunque no se ha estudiado, parece factible determinar un método automático para determinarlo analizando la imagen.



Es cierto que este método también recoge información acerca de otras texturas sólidas como puede ser el suelo, que puede resultar de interés en la descripción del resto de texturas.

Este método es muy prometedor si se consigue desarrollar las ideas expuestas anteriormente ya que funcionaría de manera automática y con todo tipo de imágenes.

### ***Algoritmo usando extracción de bordes***

Las principales ventajas de este algoritmo son:

- Gran rapidez.
- No dependiente de ningún factor como el color o la iluminación.

Los principales inconvenientes son:

- No es capaz de obtener resultados aceptables en imágenes complejas. Particularmente el algoritmo tiene problemas cuando se cruzan varias ramas.

Este método es muy interesante porque a diferencia de los anteriores no depende de las características de la imagen. Por esta razón el método a simple vista podría parecer el más prometedor, sin embargo la realidad y las pruebas no conducen hacia esa dirección..

Todos los algoritmos probados de primera y segunda derivada de extracción de bordes, no son capaces de procesar correctamente la gran complejidad que existe en la mayoría de imágenes de estas características.



# Capítulo 5

## 5. Conclusiones y trabajo futuro

### 5.1. Sobre resultados

Se han propuesto en este trabajo varias aproximaciones para tratar de resolver el problema de la identificación de los troncos de los árboles en imágenes boscosas. Algunas de estas técnicas son más precisas que otras y ninguna de ellas ofrece resultados óptimos. No obstante los resultados obtenidos se consideran suficientes para esta primera aproximación al problema. En general los métodos tienen sus ventajas e inconvenientes, por ejemplo el método usando el clasificador híbrido es bastante lento, mientras que el que usa la entropía es dependiente del nivel de iluminación en la imagen.

La principal aportación que realiza el trabajo consiste en la combinación de técnicas básicas de procesamiento de imágenes convenientemente estructuradas con el fin de lograr el proceso encaminado a la extracción de las texturas identificativas de los troncos antes mencionadas.

Los resultados obtenidos hacen que este trabajo sea una buena base de partida para posteriores experimentos en este campo.

Se encontraron varios problemas durante los estudios de los distintos métodos aplicados a la resolución del problema. Sin embargo, los principales fueron la extrema dificultad de las imágenes y la compresión de las mismas. En cuanto a la dificultad de las imágenes basta con ver la complejidad que presenta incluso para el cerebro humano. Es necesario fijarse detenidamente en las imágenes para identificar correctamente los troncos.

La compresión en las imágenes hace que se pierda gran parte de la información del color de las imágenes digitales. Ello implica que no podamos hacer uso de una de las principales armas de las que se dispone para distinguir entre troncos y ramas significativas y otros

tipos de texturas. Tal vez, disponiendo de esta información y en conjunción con alguno de los métodos propuestos se puedan optimizar los resultados.

## **5.2. Cumplimiento de los objetivos propuestos**

En relación a los objetivos generales expresados en la sección 1.2, cabe añadir que se han cumplido las expectativas previstas.

Respecto de los específicos, su grado de cumplimiento se concreta como sigue:

1. Se han estudiado una gran variedad de métodos en prácticamente la mayoría de las áreas relevantes. Esto ha permitido tener un conocimiento global bastante amplio del campo.
2. Se han utilizado imágenes estéreo de zonas boscosas
3. Se han propuesto e implementado varios métodos que dan aproximaciones más o menos precisas a la resolución del problema
4. Se proponen líneas de investigación futuras

## **5.3. Trabajos futuros**

Como trabajos futuros surgidos durante la investigación llevada a cabo se han identificado los siguientes:

- Análisis de otras representaciones del color distintas al formato RGB tales como el modelo HSI entre otros Drimbarean y Whelan (2003). Esto no pudo ser estudiado en el presente documento por la pérdida de información de color por la compresión JPEG en las imágenes disponibles.
- Desarrollar un método de correspondencia entre pares estéreo con el fin de completar la automatización del proceso, según el segundo objetivo identificado en la sección 3.2.
- Los resultados obtenidos con algunos métodos como la entropía son muy dependientes de que exista una iluminación lo más uniforme posible. Se pueden probar métodos que igualen los niveles de iluminación para obtener mejores resultados, como el filtrado homomórfico.

## 6. Bibliografía

B. Balasko, J. Abonyi and B. Feil, Fuzzy Clustering and Data Analysis Toolbox for Use with Matlab, Veszprem University, Hungary

URL:<http://www.fmt.vein.hu/softcomp/fclusttoolbox/FuzzyClusteringToolbox.pdf>

BME Texture Database (disponible on-line [http://bme.t.u-tokyo.ac.jp/tech\\_ref/texture\\_DB/search.asp](http://bme.t.u-tokyo.ac.jp/tech_ref/texture_DB/search.asp))

J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Kluwer, Plenum Press, New York, 1981.

Brodatz Texture Database (disponible on-line <http://www.ux.uis.no/~tranden/brodatz.html>)

J.C.W. Chan, N. Laporte and R.S. Defries, "Texture Classification of logged forest in tropical Africa using machine-learning algorithms," *Int. J. Remote Sensing*, vol. 24, n° 6, pp. 1401-1407, March 2003.

E.R. Davies. *Machine Vision: Theory, Algorithms, Practicalities*. Academic Press, London, 1997

A. Drimbarean and P.F. Whelan, "Experiments in colour texture analysis," *Pattern Recognition Letters*, vol. 22, n° 4, pp. 1161-1167, April 2003.

D. DeSieno, Adding a conscience to competitive learning. *Proc. IEEE Int. Conf. Neural Networks*, 1, pp. 117-124, San Diego, CA. 1988.

R.O. Duda, P.E. Hart and D.S. Stork, *Pattern Classification*, Wiley, 2000.

F. del Frate, F. Pacifici, G. Schiavon, C. Solimini, "Use of Neural Networks for Automatic Classification from High-Resolution Images," *IEEE Trans. Geoscience and Remote Sensing*, vol. 45, n° 4, pp. 800- 809, April 2007.

G. Giacinto, F. Roli and L. Bruzzone, "Combination of neural and statistical algorithms for supervised classification of remote-sensing image," *Pattern Recognition Letters*, vol. 21, no. 5, pp. 385-397, May 2000.

R.C. Gonzalez and R.E. Woods, "Digital Image Processing", Prentice-Hall, 2002

J. Grim, J. Kittler, P. Pudil and P. Somol "Multiple Classifier Fusion in Probabilistic Neural Networks," *Pattern Analysis and Applications*, 5, 221-233, 2002.

M. Guijarro, Diseño de un nuevo clasificador híbrido no supervisado para texturas naturales en imágenes. Proyecto Fin de Master en Investigación en Ingeniería Informática. Dpto. Ingeniería del Software e Inteligencia Artificial, Facultad de Informática. Universidad Complutense de Madrid, 2007.

M. Hanmandlu, V.K. Madasu and S. Vasikarla, "A Fuzzy Approach to Texture Segmentation," in *Proc. of the IEEE International Conference on Information Technology: Coding and Computing (ITCC'04)*, The Orleans, Las Vegas, Nevada, USA, April. 2004, pp. 636-642.

P.V.C. Hough, "Methods and Means for Recognizing Complex Patterns", U.S. Patent 3,069,654, 1962.

D.W. Kim, K.H. Lee and D. Lee, "Fuzzy Cluster validation index based on inter-clúster proximity," *Pattern Recognition Letters*, vol. 24, pp. 2561-2574, November 2003.

J. Kittler, M. Hatef, R.P.W. Duin and J. Matas, "On Combining Classifiers" *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226-239, March 1998.

S. Kumar, J. Ghosh and M.M. Crawford, "Hierarchical Fusion of Multiple Classifiers for Hyperspectral Data Analysis," *Pattern Analysis and Applications*, 5, 210-220, 2002.

S. Kumar, J. Ghosh and M.M. Crawford, "Best-bases feature extraction for pairwise classification of hyperspectral data," *IEEE Trans. Geoscience and Remote Sensing*, 39(7), 1368-1379, 2001.

L.I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*, Wiley, 2004.

Y. Linde, Buzo, A. and Gray, R.M, "An Algorithm for Vector Quantization Design" *IEEE Trans. Communications*, 28(1), 84-95, 1980.

S.P. Lloyd, "Least squares quantization in PCM's", *Bell Telephone Laboratories Paper*, Murray Hill, NJ, 1957.

P. Maillard, Comparing Texture Analysis Methods through Classification, *Photogrammetric Engineering and Remote Sensing*, vol. 69, nº 4, pp. 357-367, April, 2003.

A.D. Marshall and R.R. Martin, Computer Vision models and Inspection, World Scientific Series in Robotics and Automated Systems, vol. 4, World Scientific, London, 1993.

J.M. Martínez-Otzeta, Combinación de Clasificadores. En *Aprendizaje Automático: conceptos básicos y avanzados* (Serra, Ed.), Prentice-Hall, 2006.

T.S. Newman and Jain, A.K. (1995). A survey of automated visual inspection. *Computer Vision and Image Understanding*. 61(2), 231-262.

N. Otsu, N.. A threshold selection method from gray-level-histograms. *IEEE Trans. System Man and Cybernetics*. SMC-9, 62-66, 1979.

G. Pajares y J.M. de la Cruz, "Visión por Computador: Imágenes Digitales y Aplicaciones", RA-MA, 2001.

G. Pajares, V. Moreno y J. M. de la Cruz, Clasificación de texturas mediante redes neuronales, *Ingeniería Civil. Centro de Estudios y Experimentación de Obras Públicas (CEDEX).- Ministerio de Fomento*, nº 123, pp. 61-69, (2001).

G. Pajares and J. M. Cruz, Clasificación de Texturas Naturales mediante K-Means, *Revista Electrónica de Visión por Computador* <http://revc.uab.es/revista/06/>, no. 6, pp. 1-18, Ed. Centre de Visió per Computador, Universidad Autònoma Barcelona, 2002.

G. Pajares, J. M. de la Cruz y V. Moreno, Clasificación de texturas naturales mediante agrupamiento borroso, *Ingeniería Civil. Centro de Estudios y Experimentación de Obras Públicas (CEDEX).- Ministerio de Fomento*, nº 127, pp. 83-89, 2002.

D. Puig and M.A. García, "Automatic texture feature selection for image pixel classification," *Pattern Recognition*, vol. 39, nº 11, pp. 1996-2009, November 2006.

R. Rud, M. Shoshany, V. Alchanatis and Y. Cohen, "Application of spectral features' ratios for improving classification in partially calibrated hyperspectral imagery: a case study of separating Mediterranean vegetation species", *Journal Real-Time Image Processing*, vol. 1, pp. 143-152, 2006.

R.M. Valdovinos, J. S. Sánchez and R. Barandela. "Dynamic and Static weighing in classifier fusion," in *Pattern Recognition and Image Analysis*, Lecture Notes in Computer Science (J.S. Marques, N. Pérez de la Blanca and P. Pina, Eds.), Springer-Verlag, Berlin. 2005, pp. 59-66.

VisTex (disponible on-line  
<http://vismod.media.mit.edu/vismod/imagery/VisionTexture/information.html>)

H. Yu, M. Li, H.J. Zhang and J. Feng, "Color texture moments for content-based image retrieval". In Proc. International Conference on Image Processing, vol. 3, pp. 24-28, 2002.

H.J. Zimmerman, *Fuzzy Set Theory and its Applications*, Kluwer Academic Publishers, Norwell, 1991.



